

Zastosowanie głębokiego uczenia na danych hiperspektralnych do klasyfikacji materiałów przyszwyci obuwia

Paweł Wittels, Jan Jaczewski

Avicon Advanced Vision Control, Aleje Jerozolimskie 202, 02-486 Warszawa

Piotr Brzeziński

VIVE Textile Recycling Sp. z o.o., ul. Łopuszańska 22, 02-220 Warszawa

Streszczenie: Rosnąca w społeczeństwie świadomość ekologiczna sprawia, że w Europie rośnie zapotrzebowanie na używaną odzież i obuwie. Jednak wysokie koszty siły roboczej w krajach UE sprawiają, że ręczne sortowanie odzieży staje się nieopłacalne. Skuteczna klasyfikacja materiałów odzieży i obuwia w oparciu o tradycyjne rozwiązania, np. kamery mono/RGB nie jest możliwa. W artykule opisujemy opracowany i wdrożony system automatycznej klasyfikacji materiałów przyszwyci obuwia na linii sortującej, wykorzystujący sztuczne sieci neuronowe do analizy obrazu z kamer hiperspektralnych w paśmie NIR-SWIR (900–1700 nm).

Słowa kluczowe: HSI, analiza obrazu, obrazowanie hiperspektralne, sieci neuronowe, HSI-CNN

1. Wprowadzenie

Niskie koszty produkcji przemysłowej, będące rezultatem taniej siły roboczej w krajach Dalekiego Wschodu, jak również rosnącej automatyzacji produkcji, są przyczyną spadających cen nowej odzieży i obuwia w relacji do średniego wynagrodzenia. Wraz z tym trendem rośnie marnotrawstwo – coraz większe ilości odzieży i obuwia wciąż zdalnych do użytku, trafiają na wysypiska śmieci czy do spalarni. Szacuje się, że całkowita ilość odpadów tekstylnych oraz obuwia pochodzących od konsumentów może osiągnąć nawet 148 milionów ton na całym świecie do 2030 r., co stanowi wzrost o 62 % w porównaniu z 2015 r. [21].

Jednocześnie w krajach Europy rośnie świadomość ekologiczna, która sprawia, że coraz więcej osób decyduje się na zakup odzieży z drugiego obiegu nie tylko z przyczyn ekonomicznych, ale także z chęci ochrony środowiska. Oczywiście więc rozwiązaniem, oczekiwanym zarówno z przyczyn środowiskowych jak i społecznych, jest wprowadzanie używanych tekstyliów do obiegu wtórnego w modelach Circular Economy [1].

Badania wskazują, iż kwestie technologii pozwalającej na wydajne, szybkie i ekonomicznie uzasadnioną automatyzację procesów identyfikacji i sortowania używanych odpadów postkonsumenckich są jednymi z kluczowych barier rozwoju modeli Circular Economy i skutecznego zwracania używanych tekstyliów i obuwia do ponownego użycia [1].

Skuteczna identyfikacja poszczególnych obiektów (tekstylia, buty etc.) pozwala bowiem stworzyć powtarzalny proces podejmowania decyzji, czy dany obiekt ma potencjał (wartości użytkowe) do ponownego zaoferowania go klientom w ramach Circular Economy Scheme, czy też należy znaleźć dla niego inne metody utylizacji. Wyzwanie technologiczne jest szczególnie istotne dla odpadów postkonsumenckich i stanowi jedną z kluczowych barier rozwoju korzystnych dla zrównoważonego rozwoju modeli CE [2].

Powszechnie stosowana jest automatyzacja sortowania różnego typu obiektów na podstawie ich kształtu, koloru czy gabarytów. Jednakże obrazy rejestrowane przez klasyczne kamery monochromatyczne i kolorowe (RGB) zasadniczo nie pozwalają rozróżniać materiałów, z których wykonane są badane obiekty.

Rozróżnianie różnorodnych materiałów, niezależnie od ich koloru i faktury, możliwe jest natomiast dzięki zastosowaniu obrazowania hiperspektralnego. Obrazy hiperspektralne, szczególnie w zakresie bliskiej podczerwieni, przy zastosowaniu metod głębokiego uczenia maszynowego, umożliwiają efektywną analizę i rozpoznawanie materiałów na poziomie niedostępnym dla tradycyjnych metod wizyjnych.

Jednym z wyzwań w branży recyklingu i sortowania odzieży oraz obuwia używanego jest automatyczna identyfikacja materiałów, z których wykonana jest wierzchnia część obuwia. Dokładna klasyfikacja materiałów przyszwyci buta ma kluczowe znaczenie dla optymalnego zarządzania łańcuchem

Autor korespondujący:

Paweł Wittels, pawel.wittels@avicon.pl

Artykuł recenzowany

nadesłany 03.03.2025 r., przyjęty do druku 19.05.2025 r.



Zezwala się na korzystanie z artykułu na warunkach licencji Creative Commons Uznanie autorstwa 4.0 Int.

wartości, w tym dla procesów sortowania i dalszego wykorzystania lub recyklingu.

W prezentowanym artykule opisujemy projekt zrealizowany w zakładzie sortowania odzieży używanej VIVE Textile Recycling, w którym zastosowano system hiperspektralny rejestrujący obraz w zakresie obejmującym długości fali światła od bliskiej podczerwieni do podczerwieni krótkofalowej (900–1700 nm, NIR-SWIR) do automatycznej klasyfikacji materiałów obuwia. System wykorzystuje liniowe kamery hiperspektralne oraz metody głębokiego uczenia maszynowego do analizy struktury widmowej powierzchni butów. Celem było uzyskanie wysokiej skuteczności predykcji, a także stabilnej powtarzalności wyników w warunkach przemysłowych, gdzie buty są transportowane na taśmociągu i wprowadzane do ciemni pomiarowej.

1.1. Obrazowanie hiperspektralne

Obrazowanie hiperspektralne definiowane jest jako technika obrazowania, która polega na akwizycji setek zdjęć tej samej sceny, z których każde rejestrowane jest w kolejnych wąskich zakresach długości promieniowania elektromagnetycznego [3]. Typowymi zakresami kamer hiperspektralnych jest zakres UV-VIS (380–800 nm lub 400–1000 nm), NIR-SWIR (900–1700 nm), SWIR (900–2500 nm) [4] oraz MWIR (2700–5500 nm) [5]. Podział na poszczególne zakresy długości fal podyktowany jest m.in. dostępnymi technologiami sensorów (ich efektywnością kwantową w funkcji długości fali), rozdzielczościami sensorów, transmisyjnością materiałów, z których wykonywana jest optyka czy koniecznością optymalizacji konstrukcji obiektywu. Optyka stosowana w obrazowaniu hiperspektralnym musi cechować się transmisyjnością w szerokim zakresie spektrum, minimalizować zjawisko przesunięcia punktu ostrości w funkcji długości fali oraz minimalizację zjawiska aberracji chromatycznej [6]. Istotną cechą obiektywów stosowanych w obrazowaniu hiperspektralnym jest ich wysoka jasność. Jest to czynnik niezbędny dla wydajnego obrazowania, ze względu na bardzo selektywną (pod względem długości fali) rejestrację obrazu dla poszczególnych pasm. Dla kamery SPECIM FX10 (zakres obrazowania 400–1000 nm) szerokość połowkowa filtra (FWHM) wynosi 5,5 nm [7] – taka szerokość pasma przypada na pojedynczy obraz spektralny. Dla porównania, pojedynczy (zielony) kanał obrazu RGB pochodzącego z sensora Sony IMX174 charakteryzuje FWHM rzędu 125 nm [8]. Wskazuje to na konieczność stosowania wysokiej intensywności oświetlenia (ponad 20× większej dla opisywanej kamery hiperspektralnej w stosunku do RGB) oraz możliwe jasnej optyki dla uzyskania wymaganej intensywności obrazu i możliwie wysokiego SNR. Dodatkowo dla pasm w okolicach krańców zakresów pomiarowych kamer hiperspektralnych, efektywność kwantowa znacząco spada co stawia dodatkowe wyzwania przed konstrukcją oświetlenia.

Przez obraz hiperspektralny (ang. *spectral cube* [9]) rozumiemy zestaw obrazów tej samej sceny, z których każdy odpowiada oddzielnemu pasmu (definiowanemu przez FWHM oraz środkową długość fali). Każdy oddzielny obraz takiego zestawu, odpowiadający wąskiemu podzakresowi

zakresu pomiarowego, nazywany jest kanałem obrazu hiperspektralnego. Istnieje kilka strategii (technologii) uzyskiwania obrazu hiperspektralnego (rys. 1).

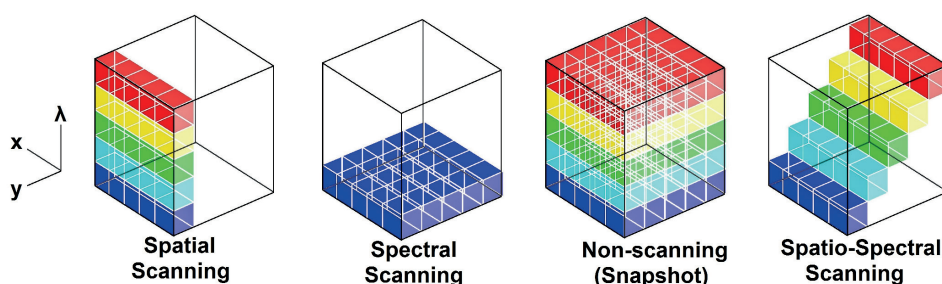
Skanowanie przestrzenne (ang. *pushbroom* [10]) zapewnia uzyskanie pełnej informacji spektralnej dla jednego wycinka (paska) obrazu. Przemieszczanie się sceny względem kamery oraz synchroniczne z tym ruchem wyzwalanie akwizycji kolejnych linii, pozwala na zebranie kompletnego obrazu. Zaletą tej technologii skanowania jest konieczność oświetlenia jedynie jednego paska sceny oraz właściwie nieograniczona rozdzielczość spektralna uzyskiwanego obrazu. Technologia ta nie wymaga unieruchamiania obiektu podczas skanowania. Skanowanie spektralne polega na zbieraniu obrazów pojedynczych pasm całej sceny. Między kolejnymi klatkami konieczna jest zmiana pasma rejestracji (przez zmianę filtrów pasmowoprzepustowych lub długości fali światła padającego na próbkę). Wymaga to unieruchomienia próbki względem kamery, prawidłowe oświetlenie całej sceny a liczba uzyskiwanych pasm jest ograniczona ze względu na technologię zmiany pasm (liczbę dostępnych filtrów czy czas przestrajania źródła światła). Obrazowanie jednorazowe, ze względu na ograniczoną rozdzielczość matrycy, wymaga kompromisu między rozdzielczością przestrzenną obrazu a liczbą rejestrowanych pasm: dla 125 pasm w zakresie 450–900 nm możliwe jest uzyskanie rozdzielczości przestrzennej 70×70 px [11], dla 16 pasm w zakresie 460–600 nm uzyskiwana jest rozdzielczość przestrzenna 512×272 px [23]. Obrazowanie tego typu wymaga oświetlenia całej analizowanej sceny ale nie wymaga unieruchamiania obiektu. Skanowanie przestrzenno-spektralne polega na pozyskaniu informacji o pojedynczych pasmach dla pojedynczych linii obrazu. Wzajemny ruch sceny i kamery pozwala na zebranie powyższych informacji dla kolejnych linii obrazu, a następnie konstrukcję obrazu hiperspektralnego. Pozwala to na uzyskanie relatywnie wysokiej rozdzielczości spektralnej i przestrzennej np. 150 pasm w zakresie 470–900 nm, rozdzielczość przestrzenna 2048 px. Konieczne jest natomiast oświetlenie przestrzenne całej sceny a nie tylko jednej linii obrazu.

Badania opisane w artykule prowadzone były z wykorzystaniem kamery Specim FX17, pracującej w technologii skanowania przestrzennego, charakteryzującą się zakresem pomiarowym 900–1700 nm, FWHM 8 nm, rozdzielczością spektralną 224 pasma oraz przestrzenną rozdzielczością liniową 640 px. Temperatura sensora stabilizowana jest przez moduł Peltiera. Do oświetlenia sceny wykorzystano zestaw lamp halogenowych, zapewniających wymaganą intensywność oraz ciągłość widmową światła w całym zakresie pomiarowym kamery.

1.2. Zastosowanie obrazowania hiperspektralnego

Techniki hiperspektralne znajdują zastosowanie w wielu dziedzinach. Są to m.in.

– recykling i gospodarka odpadami (identyfikacja surowców, segregacja materiałów) [12],



Rys. 1. Techniki obrazowania hiperspektralnego, kolejno od lewej: skanowanie przestrzenne, skanowanie spektralne, obrazowanie jednorazowe, skanowanie przestrzenno-spektralne

Fig. 1. Hyperspectral imaging techniques: spatial scanning, spectral scanning, snapshot imaging, spatio-spectral scanning [22]

- rolnictwo i przemysł spożywczy (ocena dojrzałości owoców, detekcja obić, nadgnić, przesuszeń, identyfikacja zanieczyszczeń) [13, 14],
- ochrona środowiska (monitoring jakości wody, detekcja mikro-plastiku) [15],
- przemysł wytwórczy (kontrola jakości materiałów, wykrywanie defektów) [16],
- sztuka i archeologia: analiza dzieł sztuki i zabytków w celu identyfikacji materiałów użytych przy ich tworzeniu [17].

W przypadku omawianego projektu użycie obrazowania hiperspektralnego umożliwiło automatyczną klasyfikację materiałów przyszwyc obuwia, co jest szczególnie istotne w masowej, hurtowej obróbce dużych ilości używanych butów o zróżnicowanej jakości, pochodzeniu i stanie zachowania.

Wyniki klasyfikacji materiałów wykorzystano dodatkowo do automatycznej weryfikacji prawidłowości automatycznej identyfikacji pary butów (buta lewego i buta prawego), wykonywanej za pomocą niezależnego systemu, bazującego na danych z kamer RGB. Zaobserwowano, iż co do zasady buty z tej samej pary „współdzielą” swoje cechy: proporcje udziału różnych materiałów w poszczególnych obszarach butów są zbliżone. Dzięki temu ustalenie zewnętrznym klasyfikatorem dodatkowej cechy obydwu butów (rodzaj materiału) może być jednocześnie testem prawidłowości zestawienia takiej pary (buta lewego z butem prawym) przez klasyfikator „Alina” stosowany do celów parowania obuwia. Klasyfikator bazujący na danych hiperspektralnych pozwala tym samym na ograniczenie błędów *False Positive* w procesie parowania.

1.3. Algorytmy uczenia maszynowego w analizie danych hiperspektralnych

Tradycyjne podejścia do analizy obrazowania hiperspektralnego HSI (ang. *Hyperspectral Imaging*) bazowały na metodach, takich jak k-nearest neighbors (kNN), support vector machines (SVM) czy random forests (RF) [18]. Choć te techniki pozwalały na skuteczną klasyfikację danych o niewielkiej liczbie cech, ich efektywność spadała wraz ze wzrostem liczby kanałów hiperspektralnych [19]. Dzięki rozwojowi metod głębokiego uczenia, takich jak konwolucyjne sieci neuronowe CNN (ang. *Convolutional Neural Network*), stało się możliwe efektywne przetwarzanie danych o wysokiej liczbie wymiarów bez potrzeby ich redukcji, co ograniczało utratę istotnych informacji spektralnych [20].

Metody CNN zrewolucjonizowały analizę obrazów hiperspektralnych dzięki swojej zdolności do automatycznego wyodrębniania cech spektralno-przestrzennych z danych [18]. Metody takie jak HSI-CNN wykorzystują warstwy konwolucyjne do przetwarzania danych spektralnych z zachowaniem ich przestrzennej struktury. W przypadku sieci takich jak HSI-CNN, przekształcenie lokalnych cech spektralno-przestrzennych na jednowymiarowe mapy cech umożliwia wydajną klasyfikację obrazów hiperspektralnych nawet przy ograniczonej liczbie danych treningowych [18].

W artykule [20] przedstawiono nowatorską architekturę CNN zaprojektowaną specjalnie do analizy obrazów hiperspektralnych. Model ten wprowadza trzy kluczowe innowacje:

- Głębsza i szersza sieć: W przeciwieństwie do tradycyjnych sieci CNN, Contextual CNN wykorzystuje znacznie większą liczbę warstw, co pozwala na bardziej efektywne wyodrębnianie zaawansowanych cech spektralnych i przestrzennych. Jest to osiągnięte dzięki zastosowaniu uczenia resztkowego (ang. *residual learning*), które umożliwia skuteczniejsze szkolenie głębokich sieci.
- Lokalny kontekst przestrzenno-spektralny: Model optymalnie wykorzystuje lokalne zależności między sąsiednimi pikselami, co jest kluczowe w przypadku danych hiperspektralnych, gdzie informacja przestrzenna i spektralna są ściśle

powiązane. W tym celu zastosowano bank filtrów wieloskalowych, który pozwala na jednoczesne wydobywanie cech w różnych skalach przestrzennych.

- Efektywność w ograniczonych zbiorach danych: Jednym z głównych wyzwań w analizie HSI jest niedobór danych treningowych. Dzięki zastosowaniu mechanizmu uczenia resztkowego oraz projektowaniu sieci o zoptymalizowanej liczbie parametrów, Contextual CNN osiąga doskonałe wyniki nawet przy niewielkiej liczbie przykładów treningowych.

Ze względu na ograniczoną ilość danych uczących (wynikającą przede wszystkim z wysokiej pracochłonności ręcznego maskowania obrazów uczących), a także ze względu na bardzo wysoką dokładność opisanego w [18] modelu, w opisywanym projekcie zdecydowano o użyciu modelu HSI-CNN do klasyfikacji materiałów obuwia.

1.4. Opis problemu

Celem opracowywanego Systemu NIR¹ była automatyzacja procesu klasyfikacji materiału przyszwyc butów na linii sortowania obuwia używanego. System klasyfikacji materiału jest elementem większego systemu, którego zadaniem jest nie tylko rozpoznawanie tworzywa, z którego wykonane są buty, ale również automatyczne parowanie pomieszanych butów i ich wycena. Klasyfikowane buty zawieszane są za noski na metalowych zawieszkach przenośnika podwieszanego.

2. Zaproponowane rozwiązanie

2.1. Akwizycja danych

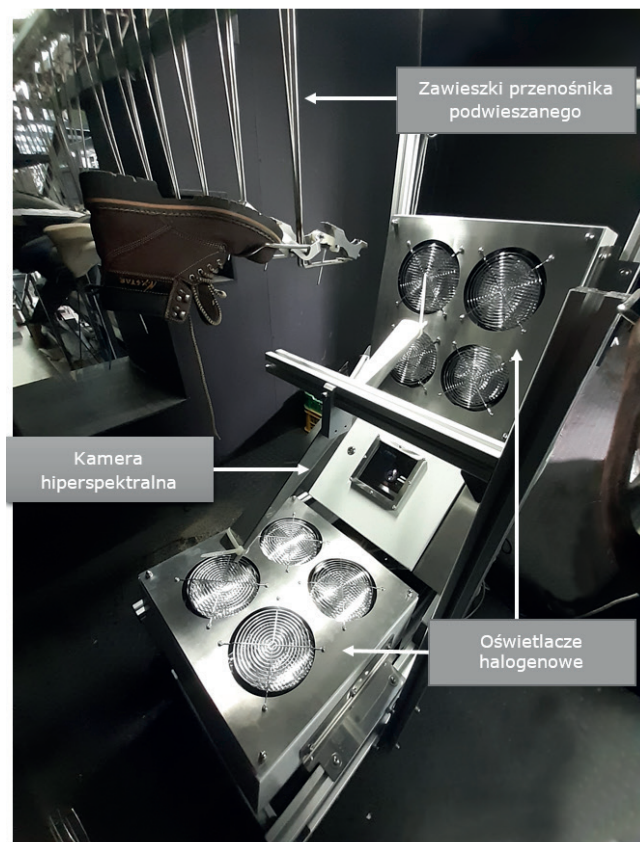
System NIR składa się z dwóch niezależnych stanowisk pomiarowych (nazwanych L1 i L2; jedno ze stanowisk przedstawiono na rys. 2), umieszczonych w celu zwiększenia przepustowości na dwóch równoległych liniach sortowania. Starano się, by stanowiska pomiarowe L1 i L2 były możliwie identyczne, jednakże ponieważ były one dostawione do dwóch istniejących już linii, dystanse i kąty obserwacji mogły się wzajemnie nieznacznie różnić. W tym celu modele trenowano na obrazach zarejestrowanych na obu systemach, z zachowaniem równych proporcji. W celu uniezależnienia wyników pomiarów od warunków oświetleniowych otoczenia, hiperspektralne kamery liniowe Specim FX-17 umieszczono wraz z odpowiednio dobranymi oświetlaczami w specjalnych ciemniach.

Rejestracja obrazu rozpoczyna się na sygnał z fotokomórki wykrywającej pojawienie się kolejnej zawieszki przejeżdżającego przez ciemnię taśmociągu. Każdorazowo rejestrowane jest 131 linii obrazu z kamery, co odpowiada szerokości zawieszki z butem z pewnym zapasem. Natychmiast po zarejestrowaniu, 131 linii łączonych jest do pojedynczego obrazu hiperspektralnego składającego się z 224 warstw (pasm) o wymiarach 640×131 px o 12-bitowej rozdzielczości tonalnej. W tych samych warunkach oświetleniowych i tych samych parametrach ekspozycji zarejestrowano również wzorec bieli, a przy braku oświetlenia również wzorec czerni, które służyły do normalizacji rejestrowanych obrazów (przestrzennej i spektralnej).

2.2. Przygotowanie danych uczących

Zarejestrowane obrazy hiperspektralne poddano redukcji liczby kanałów o 50 % w taki sposób, że do zredukowanego obrazu hiperspektralnego przepisano co drugi kanał obrazu

1 Nazwa Systemu NIR nawiązuje do zakresu pomiarowego zastosowanej kamery hiperspektralnej FX-17 (900–1700 nm) – klasyfikowanej przez jej producenta, firmę Specim, jako bliską podczerwień (ang. Near-Infrared – NIR), mimo iż w rzeczywistości pokrywa ona również część zakresu podczerwieni krótkofalowej (ang. Short-Wave Infrared – SWIR)



Rys. 2. System NIR – umieszczona w ciemni kamera hiperspektralna rejestruje obraz butów mocowanych za noski na zawieszkach przenośnika podwieszanego. Zestawy oświetlaczy halogenowych doświetlają scenę z dwóch kierunków
Fig. 2. NIR System – a hyperspectral camera placed in a darkroom records an image of shoes mounted by their toes on the overhead conveyor. Halogen light sets illuminate the scene from two directions

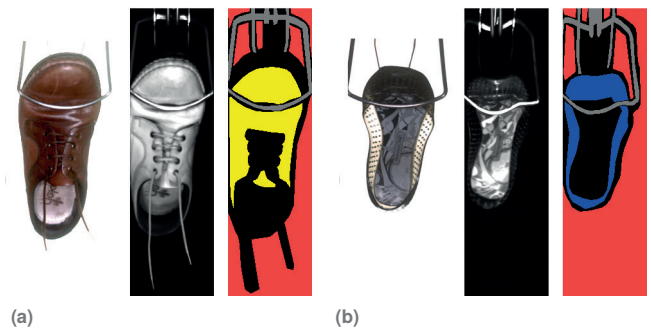
oryginalnego, a nowa wartość piksela na każdym z kanałów wyliczana była jako średnia z wartości trzech sąsiadujących kanałów obrazu oryginalnego: danego, poprzedzającego oraz następującego. Etap ten dodano po pierwszych testach, w celu skrócenia czasu predykcji finalnego modelu, przy pomijalnym pogorszeniu precyzji. W finalnym rozwiązaniu redukcja liczby kanałów wykonywana była już w kamerze ($2\times$ binning spektralny). Mechanizm *binningu* w kamerze jest analogiczny do opisanego powyżej, zastosowanego na danych uczących.

Zarejestrowany dataset, przeznaczony zarówno do uczenia sieci, jak i jej późniejszego testowania i oceny, zawierał 6616 obrazów hiperspektralnych. Zarejestrowane dane obejmowały obrazy 615 butów oraz 47 pustych zawieszek zarejestrowanych

przez oba punkty pomiarowe (L1 i L2). Każdy but zarejestrowany był kilkakrotnie na każdym stanowisku. Pliki obrazów zawierały unikalny numer ID obrazowanego buta, co pozwoliło na etapie przygotowania zbioru grupować obrazy tego samego egzemplarza obuwia.

Spośród 615 butów do zbioru uczącego (trenującego i walidacyjnego) wytypowano 230 butów: 63 ze skóry, 54 z gumy, 54 z tworzywa sztucznych i 59 z tekstyliów naturalnych. Buty do poszczególnych klas zbioru uczącego nie były wybierane całkiem losowo – starano się pilnować zróżnicowania pod względem ich koloru, jasności oraz podklas materiału. Do zbioru uczącego wybrano po dwa zdjęcia hiperspektralne każdego buta (po jednym wykonanym na każdym ze stanowisk), tzn. 460 obrazów hiperspektralnych. Obrazy pozostałych 385 butów przeznaczono do testowania algorytmu i oceny jakości klasyfikacji.

Następnie obrazy przeznaczone do uczenia sieci poddano ręcznemu maskowaniu (rys. 3). Ponieważ wejściem wybranego modelu sieci są *patches* (trójwymiarowe wycinki obrazu hiperspektralnego) o wymiarach $3 \times 3 \times 112$, a nie całe obrazy, maskowaniu poddawano jedynie obszary, których materiał nie budził żadnych wątpliwości. W szczególności wystrzegano się maskowania zapiek, naszywek, widocznych fragmentów podeszw i wnętrza butów i innych elementów stanowiących dodatki



Rys. 3. Przykłady przygotowania masek butów ze skóry (a) i z gumy (b). Obrazy każdego z butów przedstawiają odpowiednio: zdjęcie poglądowe z kamery wizyjnej (po lewej), wybrany kanał z obrazowania hiperspektralnego (w środku) i ręcznie przygotowaną maskę (po prawej stronie). Maskowano jedynie obszary o wysokiej pewności. Czarne obszary masek nie były brane pod uwagę podczas uczenia modelu. (Widoczne zniekształcenie obrazu hiperspektralnego (b) wynika z poruszenia zawieszki podczas rejestracji obrazu kamerą liniową)

Fig. 3. Examples of masks prepared for leather shoes (a) and rubber shoes (b). Images of each shoe show: an overview image from a vision camera (on the left), the selected channel from hyperspectral imaging (in the middle), and the manually prepared mask (on the right). Only areas with high confidence were masked. Black areas of the masks were not considered during model training. (The visible distortion in the hyperspectral image (b) is caused by movement of the hanger during image registration by the line scan camera)

Tab. 1. Podział na klasy i podklasy materiałów
Tab. 1. Classification of materials into main classes and subclasses

Klasa	Podklasy	Liczba butów (obrazów) w zbiorze uczącym
Skóra	zamsz, nubuk, skóra licowa	63 (126)
Guma	guma	54 (108)
Tworzywo sztuczne	ekoskóra, nylon	54 (108)
Tekstylia	bawełna/len	59 (118)
Zawieszka	klasa pomocnicza do wykrywania metalowej zawieszki	(47)
Tło	klasa pomocnicza do wykrywania tła	(507)

z niepewnego materiału. Pomijano również piksele leżące na krawędzi dwóch materiałów (będące w dużej mierze uśrednieniem pomiaru każdego z materiałów).

Obrazy hiperspektralne zostały następnie poddane normalizacji względem obrazów referencyjnej bieli i czerni. Normalizacja względem czerni pozwala zredukować wpływ szumów na matrycy, a względem bieli – zminimalizować wpływ niejednorodności oświetlenia zarówno w zakresie spektralnym jak i przestrzennym.

Następnie maski obrazów wybranych do zbioru uczącego (treningowego i walidacyjnego) przekształcono w zbiór *patchy*, dla każdego piksela maski, dla którego wszystkie sąsiadujące 8 pikseli należało do maski tej samej klasy, pobierano z obrazu hiperspektralnego odpowiadający mu (mający te same współrzędne x, y) *patch* o wymiarach $3 \times 3 \times 112$ px. *Patche* zapisywano w sześciu niezależnych zbiorach poszczególnych klas (tab. 1). Każdy z tak przygotowanych zbiorów dzielony był na dwa podzbiory, zawierające odpowiednio *patche* treningowe oraz walidacyjne.

2.3. Model sieci

Jako klasyfikator wykorzystano autorski model HSI-CNN o architekturze wzorowanej na opisanej w artykule [18].

Model HSI-CNN został zaprojektowany do klasyfikacji danych hiperspektralnych przez jednoczesną analizę cech spektralnych i przestrzennych. Wejściem modelu opracowanego w ramach projektu są *patche* o wymiarach $3 \times 3 \times 112$ px, gdzie pierwsze dwa wymiary reprezentują lokalny obszar przestrzenny, a ostatni wymiar odnosi się do liczby pasm spektralnych. Proces przetwarzania rozpoczyna się od trójwymiarowej warstwy konwulucyjnej, która zawiera 128 filtrów o rozmiarze $3 \times 3 \times 24$ px i funkcji aktywacji ReLU. Warstwa ta analizuje jednocześnie strukturę przestrzenną oraz lokalne zależności spektralne. W efekcie powstaje tensor o wymiarach $1 \times 1 \times 10 \times 128$ px, który następnie zostaje przekształcony do postaci dwuwymiarowej za pomocą operacji Reshape, zmieniając jego format na $10 \times 128 \times 1$ px.

Po redukcji wymiarów spektralnych model przechodzi do analizy przestrzennej za pomocą warstwy konwulucyjnej 2D, w której zastosowano 64 filtry o wymiarach 3×3 i aktywację ReLU. Ta operacja umożliwia ekstrakcję lokalnych cech przestrzennych, prowadząc do uzyskania tensora o wymiarach $8 \times 126 \times 64$. W celu dalszej redukcji wymiarów oraz zwiększenia efektywności obliczeniowej zastosowano warstwę MaxPooling z oknem 2×2 , co powoduje zmniejszenie rozmiaru tensora do $4 \times 63 \times 64$. Następnie dane są przekształcane do jednowymiarowego wektora o długości 16 128, co przygotowuje je do klasyfikacji w pełni połączonymi warstwami neuronowymi.

Pierwsza warstwa w pełni połączona zawiera 1024 neurony z aktywacją ReLU, co pozwala na dalszą ekstrakcję cech istotnych dla klasyfikacji. Aby zapobiec przeuczeniu modelu, zastosowano mechanizm *Dropout* (losowe wyłączanie neuronów) z parametrem 0,5, który losowo deaktywuje połowę neuronów podczas treningu. Kolejna warstwa w pełni połączona składa się ze 100 neuronów, również z aktywacją ReLU, po której ponownie zastosowano warstwę *Dropout* o wartości 0,5. Ostatecznie model kończy się warstwą wyjściową o sześciu neuronach z aktywacją Softmax, której zadaniem jest przypisanie prawdopodobieństwa do jednej z sześciu klas.

Architektura modelu obejmowała dziewięć warstw z parametrami, w tym dwie warstwy splotowe z łączną liczbą 192 filtrów oraz trzy w pełni połączone warstwy. Liczba trenowanych parametrów wyniosła 16,6 mln. Model uczono od podstaw – nie stosowano transferu wag ani wcześniejszego uczenia transferowego. Podczas treningu jako funkcję straty zastosowano ważoną kategoriową entropię krzyżową (ang. *weighted categorical cross-entropy*), aby uwzględnić nierównomierny rozkład klas w zbiorze danych. Do optymalizacji zastosowano algorytm RMSprop. Proces uczenia prowadzono z wielkością partii równą 1000 oraz z maksymalną liczbą 100 epok, choć rzeczywista liczba epok

była determinowana przez mechanizm wczesnego zatrzymywania (ang. *Early Stopping*), który przerywał trening, gdy strata walidacyjna nie poprawiała się przez dziesięć kolejnych epok. Aby ograniczyć ryzyko przeuczenia, w architekturze zastosowano warstwę *Dropout* z prawdopodobieństwem 0,5 oraz strategię redukcji kroku uczenia, która zmniejszała wartość współczynnika uczenia, gdy strata walidacyjna przestawała się obniżać.

HSI-CNN stanowi hybrydowe podejście, w którym ekstrakcja cech spektralnych realizowana jest przez konwolucję 3D, natomiast analiza zależności przestrzennych odbywa się przy pomocy klasycznej konwulucji 2D. Dzięki zastosowaniu głębokiej architektury i warstw *Dropout* model jest mniej podatny na przeuczenie, a jednocześnie zdolny do skutecznej klasyfikacji danych hiperspektralnych, uwzględniając zarówno ich komponenty spektralne, jak i przestrzenne.

Wynikiem działania klasyfikatora jest obraz składający się z sześciu kanałów. Każdy piksel obrazu jest 6-elementowym wektorem, w którym każdy element ma wartość prawdopodobieństwa przynależności piksela do jednej z sześciu klas (cztery klasy materiału butów, tła oraz zawieszki). Wartości te sumują się do 100 %. Zbiorem pikseli, których klasyfikacja wliczana jest do finalnych wyników predykcji są piksele obrazu znajdujące się pod maską buta. Maska buta jest zbiorem wszystkich pikseli obrazu, pomniejszonym o piksele, które prawdopodobnie są strzemiem lub tłem (tzn. z najwyższym prawdopodobieństwem przynależą do jednej z tych dwóch klas). Maska dodatkowo jest pomniejszana o piksele, których dwa najwyższe prawdopodobieństwa różnią się wzajemnie o mniej niż pięć punktów procentowych. Na koniec maska poddana jest erozji morfologicznej, żeby pozbyć się pikseli znajdujących się na granicy buta i tła, które mogą być klasyfikowane w sposób nieprzewidywalny. Do ostatecznego wyniku predykcji klasyfikacji buta liczona jest średnia arytmetyczna prawdopodobieństw z czterech klas materiałów butów dla pikseli znajdujących się pod maską. Dopiero uzyskane globalne wyniki zostają poddane proporcjonalnej normalizacji tak, by (po usunięciu współczynników przynależności do tła lub zawieszki) sumowały się do 100 %.

3. Rezultaty

3.1. Dokładność predykcji

Średnia dokładność predykcji na zbiorze testowym 1897 obrazów hiperspektralnych 385 różnych butów zarejestrowanych na obu stanowiskach NIR wyniosła 81,23 %. Szczegółową macierz pomyłek przedstawiono na rys. 4(a). Szczególnie niskie wartości dokładności uzyskano dla obuwiu, dla którego głównym deklarowanym materiałem były tworzywa sztuczne (ekoskóra i nylon): 67,3 %. Należy jednak zwrócić uwagę, że wysoki odsetek predykcji butów z tworzywa sztucznego wskazywał na skórę naturalną (aż 17,1 %). Jest to znamienne, ponieważ charakterystyka spektralna skóry naturalnej różni się znacząco od charakterystyki tworzyw sztucznych. Widoczne jest to w przypadku obuwiu z gumy, które w prawie 15 % przypadków mylone były z butami z tworzywa sztucznego, ale w żadnym przypadku ze skórą naturalną. Wartości te nasuwają podejrzenie, że mogło dojść do błędów w przygotowaniu zbiorów uczących i walidacyjnych przez pracowników sortowni. Dlatego prawdopodobnie skóra naturalna przez klasyfikator najczęściej była „mylona” z tworzywami sztucznymi (11,2 %), a dużo rzadziej z tekstyliami naturalnymi (3,7 %), pomimo większego podobieństwa w spektrum, a guma, która dla człowieka w żadnym wypadku nie przypomina skóry ani tekstyliów, mylona była tylko z tworzywami sztucznymi (których prawdopodobnie część butów gumowych ma domieszki), a nigdy z naturalnymi tekstyliami czy skórą. Drugi czynnik, który mógł również wpłynąć na wysoką liczbę pomyłek pomiędzy skórą naturalną i ekologiczną, to potencjalne stosowanie past do butów w przy-

padku pewnej liczby butów zarówno ze skóry naturalnej jak i ekologicznej. Aspekt ten jednak na razie nie został rozpoznany i wymaga przeprowadzenia oddzielnych badań.

Otrzymane wyniki pozwoliły sformułować szereg zaleceń niezbędnych do sformułowania warunków koniecznych do uzyskania klasyfikatorów o skokowo wyższej jakości predykcji (wnioski opisane w punkcie 4 i częściowo wdrożone w dalszych pracach).

3.2. Precyzja i zgodność predykcji L1-L2

Z punktu widzenia potrzeb klienta ważniejszym parametrem od poprawności predykcji, była jej zgodność dla systemów L1 i L2. Wynika to z funkcjonalności systemu, w którym ważniejsze jest znalezienie poprawnej pary butów (z której każdy może być klasyfikowany na innym stanowisku) niż rozpoznanie ich tworzywa. Opracowany system wykazał się bardzo wysoką powtarzalnością pomiarów w ramach poszczególnych stanowisk pomiarowych. Powtarzalność predykcji klasy całego buta w ramach NIR-L1 wyniosła 99,58 %, a w przypadku NIR-L2: 99,67 %.

	Leather	Rubber	Plastic	Textile
Leather	84,5 %	0,5 %	11,2 %	3,7 %
Rubber	0,0 %	85,1 %	14,9 %	0,0 %
Plastic	17,1 %	9,7 %	67,3 %	5,9 %
Textile	4,2 %	6,6 %	1,3 %	87,8 %

(a)

L1 L2	Leather	Rubber	Plastic	Textile
Leather	94,4 %	0,0 %	4,9 %	0,6 %
Rubber	0,0 %	98,6 %	0,7 %	0,7 %
Plastic	4,1 %	3,4 %	91,2 %	1,4 %
Textile	2,4 %	2,4 %	0,6 %	94,5 %

(b)

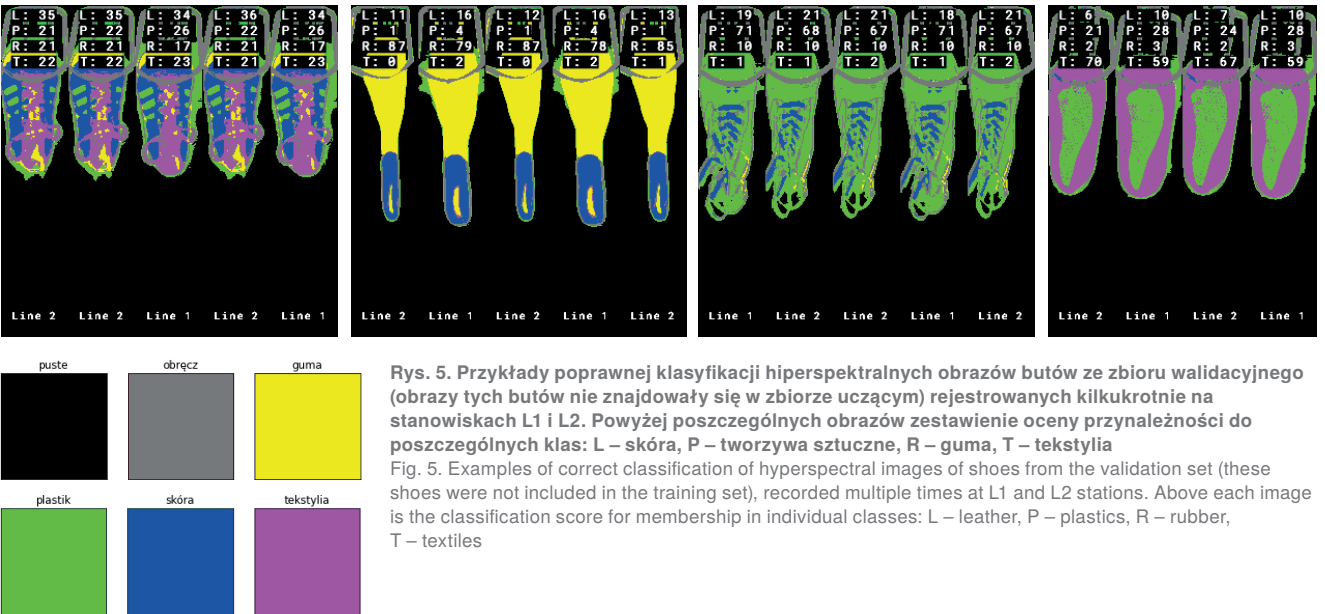
Rys. 4. Macierz pomyłek opracowanego modelu na zbiorze testowym (a) oraz macierz zgodności predykcji L1-L2 (b)
Fig. 4. Confusion matrix of the developed model on the test set (a) and prediction consistency matrix for L1-L2 (b)

Udało się również uzyskać dość wysoką zgodność predykcji tych samych butów na obu stanowiskach. Średnia zgodność predykcji L1-L2 wyniosła 94,63 %. Szczegółową macierz zgodności predykcji przedstawiono na rys. 4 (b).

Mniejsza zgodność klasyfikacji między dwoma stanowiskami, niż powtarzalność w ramach jednego stanowiska, wynika z pewnością z relacji geometrycznych układu oświetlacze-kamera-obiekt (przenośnik). Stanowiska NIR nie były kompaktowymi urządzeniami, a konstrukcjami dobudowanymi na istniejących liniach. Mimo starań, idealne zamontowanie poszczególnych elementów systemu w dwóch różnych miejscach w zakładzie sortującym nie było możliwe. Kąty i dystanse padania światła oraz kąty i dystanse obserwacji między stanowiskami różniły się nieznacznie, co mogło wpływać zarówno na różnice w intensywności pikseli na obrazie, jak i na zmianę perspektywy obserwacji obiektu, co nieraz mogło zaburzyć proporcje obrazu powierzchni wnętrza buta do powierzchni obrazu części zewnętrznej. W przypadku butów, dla których

Tab. 2. Przykłady błędnej predykcji
Tab. 2. Examples of incorrect predictions

Klasa	Wynik predykcji	Zdjęcie poglądowe
skóra		
ekoskóra		



wyniki predykcji co najmniej dwóch dominujących klas były bardzo zbliżone, skutkować to mogło odmienną finalną klasyfikacją do jednej kategorii (tab. 2, drugi wiersz).

3.3. Omówienie przykładów

Zdecydowana większość butów była klasyfikowana poprawnie, zgodnie z klasyfikacją referencyjną określoną przez specjalistów. Przykłady poprawnej klasyfikacji przedstawiono na rys. 5.

Czarny, połyskliwy, skórzany but został zaklasyfikowany w znacznej mierze jako wykonany z tekstyliów naturalnych, choć w rzeczywistości prawdopodobnie tylko sznurowadła i szwy były bawelniane. Przyczyn błędnej klasyfikacji może być kilka:

1. buty czarne były niedoreprezentowane w procesie uczenia;
2. błyszcząca powierzchnia pod różnymi kątami mogła dawać odpowiedź o zupełnie różnej intensywności;
3. potencjalnie na odpowiedź spektralną wpływ mogła mieć również pasta do butów.

Klasyfikatory na obu liniach zaklasyfikowały ten but niemal z równym prawdopodobieństwem do wszystkich czterech klas. Nosek tego buta jednoznacznie zaklasyfikowany został do skóry naturalnej, boki przyszwę – do skóry sztucznej, sznurowadła do naturalnych tekstyliów, a wnętrze buta jako guma lub tekstylia. Wnętrze buta stanowi dużą trudność w klasyfikacji – niewielka różnica kąta padania światła powoduje powstawanie cieni – różnice w intensywności pikseli we wnętrzu buta.

4. Podsumowanie i wnioski

Uzyskane rezultaty projektu należy uznać za niedoskonałe, ale jednocześnie bardzo obiecujące, charakteryzujące się wysoką powtarzalnością pomiaru i stosunkowo dobrą predykcją, szczególnie w niektórych klasach materiałowych. Analiza przeprowadzona po zakończeniu projektu pozwoliła zidentyfikować szereg czynników, których poprawa pozwoliłaby z dużym prawdopodobieństwem znacznie poprawić skuteczność klasyfikacji:

- **wielkość zbioru** – do wytrenowania wysokoskutecznych sieci neuronowych potrzeba bardzo dużej liczby różnorodnych i reprezentatywnych danych uczących. Tym bardziej, że różnorodność materiałów, tworzyw i barwników stosowanych w przemyśle obuwniczym jest bardzo duża. Wyzwanie stanowi pracochłonność, a zatem wysokie koszty ręcznego, skrupulatnego maskowania obrazów uczących. W ramach zrealizowanego projektu model sieci uczono na 460 obrazach 230 różnych butów, po ok. 50 z każdej klasy materiałowej. Szacujemy, że do uzyskania znacznie wyższej skuteczności klasyfikacji potrzeba zbioru co najmniej o rząd wielkości większego.
- **wiarygodność danych referencyjnych** – kolejnym warunkiem koniecznym do uzyskania wysokowydajnego klasyfikatora jest wykorzystywanie do uczenia danych o wysokiej wiarygodności. W przypadku zrealizowanego projektu obiektami, na których uczono modele, były buty używane. Informacja o materiale pochodziła nie od producenta obuwia, lecz od pracowników zajmujących się ręcznym sortowaniem obuwia. W toku projektu wyeliminowano zidentyfikowane pomyłki w przygotowanych danych referencyjnych, nie wiadomo jednak, ile niepoprawnie opisanych butów pozostało w danych uczących i walidacyjnych. Docelowy model powinien być uczony na informacjach od producenta lub ustalonych laboratoryjnie.
- **wpływ impregnatów i past** – zupełnie niezbadanym pozostaje aspekt potencjalnego wpływu na charakterystykę spektralną i wyniki klasyfikacji past do butów i impregnatów, które z pewnością występują na części butów z rynku wtórnego. Z dużym prawdopodobieństwem mają one znaczny wpływ na odpowiedź spektralną obrazowanych materiałów. Wpływ ten należałoby zidentyfikować i bądź to wydzielić

do oddzielnej podklasy, bądź to świadomie włączać do zbiorów uczących i walidacyjnych zarówno buty pokryte takimi środkami jak i od nich wolne – w odpowiednich proporcjach.

- **identyfikacja elementów składowych obuwia** – główna klasa obuwia (buty skórzane, tekstylne itd.) mówi jedynie o dominującym materiale przyszwy obuwia. Tymczasem buty składają się zwykle z wielu innych elementów: podeszwy, wnętrza, sznurowadeł/zamków, naszywek i in. Zdarza się również, że różne elementy przyszwę wytworzone są z innych elementów (np. gumowe noski w tenisówkach). Stanowi to wyzwanie przy ręcznym maskowaniu (należy bardzo uważnie zidentyfikować, które obszary faktycznie wykonano z „głównego” materiału), zwłaszcza, że „annotatorzy” maskujący obrazy ze względów logistycznych nie mają fizycznego dostępu do butów, a jedynie do ich obrazów. Główny problem dotyczy jednak ostatecznej klasyfikacji – w omawianym projekcie ostateczna predykcja materiału buta odnosiła się do klasy o największej powierzchni w obszarze buta. Tymczasem w przypadku niektórych butów (np. niskich butów, takich jak na rys. 3b) obraz wnętrza buta ma znacznie większą powierzchnię niż obraz ich przyszwę. Rozwiązaniem problemu mogłoby być użycie dodatkowej sieci neuronowej, która operując zarówno na obrazie hiperspektralnym, jak i na wynikach klasyfikacji z naszego modelu, identyfikowałaby obszary przyszwę, a pomijała obszary podeszwy, sznurowadeł i mniej istotnych elementów.
- **analizy laboratoryjne składu tworzyw** – wytrenowany klasyfikator często mylił klasę „guma” z klasą „tworzywa sztuczne”. Zachodzi podejrzenie, że obuwiu określane jako gumowe może być w rzeczywistości wytworzone zarówno z gumy naturalnej, gumy syntetycznej, ich mieszanek – jak i mieszanek gumy z polimerami (plastomerami lub elastomerami). Aby poprawić klasyfikację obuwia przypisywanego obecnie do klasy „guma” należy przeprowadzić badania rzeczywistego składu różnych tworzyw i następnie, w zależności od wyników analiz, rozważyć ich podział na węższe klasy (np. guma syntetyczna, guma naturalna) albo przynajmniej dobrać wyczerpujące, reprezentatywne i równoliczne przykłady tych tworzyw do zbioru uczącego.

W ramach dalszych prac planowany jest rozwój opracowanej metody przez wdrożenie rozwiązań wynikających z powyższych wniosków, jak również wykorzystanie naszej metody do klasyfikacji materiałów tekstylnych odzieży używanej.

Pierwsze dwa wnioski (zwiększenie wielkości zbioru z precyzyjniejszym określeniem danych referencyjnych) wydają się oczywiste i skala ich wpływu na finalne rozwiązanie powinna być znaczna. Jest to naturalny kierunek zwiększania precyzji sieci neuronowych i szacujemy, iż powinien doprowadzić do uzyskania dokładności na poziomie $> 95\%$.

Bardzo interesujący jest wpływ innych elementów składowych buta, widocznych na zdjęciach, a nie stanowiących przyszwę, na wnioskowanie klasyfikatora o materiale dominującym. Został on wytypowany do dalszych prac eksperymentalnych przy tworzeniu kolejnej generacji klasyfikatora. Zwiększenie rozdzielczości identyfikowania z sześciu części obrazu buta na $\pm 4 \text{ mm}^2$ pola powinien pozwolić połączyć sieci odpowiedzialne za identyfikację materiału z sieciami odpowiadającymi za interpretację obrazu. Powinno to pozwolić na stworzenie rozwiązania pozwalającego określić materiał z jakiego obuwia jest wykonane oddzielnie dla składowych części obuwia i wyeliminować jedno z istotniejszych źródeł błędów obecnego rozwiązania.

Podziękowania

Projekt POIR.01.01.01-00-0532/19 „Automatyzacja procesu identyfikacji obuwia z wykorzystaniem nowoczesnych metod uczenia maszynowego” współfinansowany przez Unię Europej-

ską ze środków Europejskiego Funduszu Rozwoju Regionalnego w ramach Programu Inteligentny Rozwój. Projekt realizowany w ramach konkursu Narodowego Centrum Badań i Rozwoju: Szybka Ścieżka.

Bibliografia

1. Saccani N., Bressanelli G., Visintin F., *Circular supply chain orchestration to overcome Circular Economy challenges: An empirical investigation in the textile and fashion industries*, "Sustainable Production and Consumption", Vol. 35, 2023, 469–482, DOI: 10.1016/j.spc.2022.11.020.
2. Wójcik-Karpacz A., Karpacz J., Brzeziński P., Pietruszka-Ortyl A., Ziębicki B., *Barriers and Drivers for Changes in Circular Business Models in a Textile Recycling Sector: Results of Qualitative Empirical Research*, "Energies", Vol. 16, No. 1, 2023, DOI: 10.3390/en16010490.
3. Yuen P. W.T., Richardson M., *An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition*, "The Imaging Science Journal", Vol. 58, No. 5, 2010, 241–253, DOI: 10.1179/174313110X12771950995716.
4. ElMasry G., Sun D.-W., *CHAPTER 1 – Principles of Hyperspectral Imaging Technology*, "Hyperspectral Imaging for Food Quality Analysis and Control", 2010, 3–43, DOI: 10.1016/B978-0-12-374753-2.10001-2.
5. Daveri A., Paziani S., Marmion M., Harju H., Vidman A., Azzarelli M., Vagnini M., *New perspectives in the non-invasive, in situ identification of painting materials: The advanced MWIR hyperspectral imaging*, "TrAC Trends in Analytical Chemistry", Vol. 98, 2018, 143–148, DOI: 10.1016/j.trac.2017.11.004.
6. Liu Y., Yang Z.-H., Yu Y.-J., Wu L.-A., Song M.-Y., Zhao Z.-H., *Chromatic-Aberration-Corrected Hyperspectral Single-Pixel Imaging*, "Photonics", Vol. 10, No. 1, 2023, DOI: 10.3390/photonics10010007.
7. Ye Seong Kang, Chan Seok Ryu, Jeong Gyun Kang, *Presenting a Multispectral Image Sensor for Quantification of Total Polyphenols in Low-Temperature Stressed Tomato Seedlings Using Hyperspectral Imaging*, "Sensors", Vol. 24, No. 13, 2024, DOI: 10.3390/s24134260.
8. Skauli T., *A Performance Characteristic for Net Light Collection in Hyperspectral and Conventional Cameras: A **, "IEEE Transactions on Geoscience and Remote Sensing", Vol. 60, 2022, DOI: 10.1109/TGRS.2022.3228071.
9. Qingli Li i in., *Review of spectral imaging technology in biomedical engineering: achievements and challenges*, "Journal of Biomedical Optics", Vol. 18, No. 10, 2013, DOI: 10.1117/1.JBO.18.10.100901.
10. Boldrini B., Kessler W., Rebner K., Kessler R.W., *Hyperspectral Imaging: A Review of Best Practice, Performance and Pitfalls for in-line and on-line Applications*, "Journal of Near Infrared Spectroscopy", Vol. 20, No. 5, 2012, 483–508, DOI: 10.1255/jnirs.1003.
11. Sousa J.J., i in., *UAV-Based Hyperspectral Monitoring Using Push-Broom and Snapshot Sensors: A Multisite Assessment for Precision Viticulture Applications*, "Sensors", Vol. 22, No. 17, 2022, DOI: 10.3390/s22176574.
12. Palmieri R., Gasbarrone R., Fiore L., *Hyperspectral Imaging for Sustainable Waste Recycling*, "Sustainability", Vol. 15, No. 10, 2023, DOI: 10.3390/su15107752.
13. Saha D., Manickavasagan A., *Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review*, "Current Research in Food Science", Vol. 4, 2021, 28–44, DOI: 10.1016/j.crfs.2021.01.002.
14. Mehrad Nikzadfar i in. *Hyperspectral Imaging Aiding Artificial Intelligence: A Reliable Approach for Food Qualification and Safety*, "Applied Sciences", Vol. 14, No. 21, 2024, DOI: 10.3390/app14219821.
15. Stuart M.B., McGonigle A.J.S., Willmott J.R., *Hyperspectral Imaging in Environmental Monitoring: A Review of Recent Developments and Technological Advances in Compact Field Deployable Systems*, "Sensors", Vol. 19, No. 14, 2019, DOI: 10.3390/s19143071.
16. Yan Y., Ren J., Sun H., *Nondestructive Quality Control in Powder Metallurgy using Hyperspectral Imaging*, "Computer Vision and Pattern Recognition", 2022, DOI: 10.48550/arXiv.2207.12966.
17. Gao Z., Du M., Cao N., Hou M., Wang W., Lyu S., *Application of hyperspectral imaging technology to digitally protect murals in the Qutan temple*, "Heritage Science", Vol. 11, No. 1, 2023, DOI: 10.1186/s40494-022-00847-7.
18. Luo Y., Zoi J., Yao C., Zhao X., Li T., Bai G., *HSI-CNN: A Novel Convolution Neural Network for Hyperspectral Image*, 2018 International Conference on Audio, Language and Image Processing (ICALIP), 2018, 464–469, DOI: 10.1109/ICALIP.2018.8455251.
19. Khan M.J., Khan H.S., Yousaf A., Khurshid K., Abbas A., *Modern Trends in Hyperspectral Image Analysis: A Review*, "IEEE Access", Vol. 6, 2018, 14118–14129, DOI: 10.1109/ACCESS.2018.2812999.
20. Lee H., Kwon H., *Going Deeper With Contextual CNN for Hyperspectral Image Classification*, "IEEE Transactions on Image Processing", Vol. 26, No. 10, 2017, 4843–4855, DOI: 10.1109/TIP.2017.2725580.

Inne źródła

21. European Environment Agency. EEA Report No 15/2021. European Environment Agency Report, 2021.
22. [https://ca.m.wikipedia.org/wiki/Fitxer:AcquisitionTechniques.jpg].
23. XIMEA GmbH, *xiSPEC2 – hyperspectral imaging camera series – technical documentation*, [www.ximea.com/getattachment/402901d5-1fa3-4d27-b56e-bafb5f33156e/xiSpec_TechnicalManual-DWL_manual.pdf].

Application of Deep Learning on Hyperspectral Data for Classification of Shoe Upper Materials

Abstract: The growing environmental awareness in society is driving an increasing demand for second-hand clothing and footwear in Europe. However, the high labor costs in EU countries make manual sorting of clothing not economically viable. Effective classification of garment and footwear materials using traditional solutions, such as mono/RGB cameras, is not possible. In our article, we describe the development and implementation of an automated system for classifying upper shoe materials on a sorting line, utilizing artificial neural networks to analyze images from hyperspectral cameras in the NIR-SWIR range (900–1700 nm).

Keywords: HSI, image analysis, hyperspectral imaging, neural networks, HSI-CNN

mgr inż. Paweł Wittels

pawel.wittels@avicon.pl
ORCID: 0000-0001-7630-6629

Absolwent kierunku Automatyka i Robotyka na Wydziale Inżynierii Produkcji Politechniki Warszawskiej (2011). Pracował m.in. w Przemysłowym Instytucie Automatyki i Pomiarów PIAP oraz PIAP Space. Zajmuje się komputerową analizą obrazu, w tym obrazowaniem hiperspektralnym. Autor artykułów i prelekcji na międzynarodowych konferencjach naukowych. Obecnie pełni rolę kierownika ds. rozwoju oprogramowania w firmie Avicon Advanced Vision Control.



mgr inż. Jan Jaczewski

info@avicon.pl
ORCID: 0009-0007-4534-3161

Absolwent kierunku Automatyka i Robotyka na Wydziale Mechanicznym Energetyki i Lotnictwa Politechniki Warszawskiej. Specjalizuje się w tematyce wizji maszynowej, ze szczególnym uwzględnieniem wykorzystania algorytmów uczenia maszynowego do analizy i klasyfikacji obrazów w czasie rzeczywistym. Kierownik projektów badawczo-rozwojowych dotyczących obrazowania hiperspektralnego i przetwarzania danych z niego pochodzących. Autor i współautor patentów polskich i europejskich



mgr Piotr Brzeziński

piotr.brzezinski@vive.com.pl
ORCID: 0000-0001-9487-5470

Dyrektor Biura Badań i Rozwoju w VIVE Textile Recycling. Członek Rady Naukowej – Sieć Badawcza Łukasiewicz – Instytut Chemii Przemysłowej. Wykładowca akademicki, ekspert z bardzo dużym doświadczeniem z obszaru zarządzania projektami B+R, tworzenia ekosystemu B+R, inwestycji w sferę wysokich technologii. Od wielu lat zajmuje się tematyką ewaluacji jednostek naukowych, czasopism naukowych, badaniami naukowymi. Współtworzył system oceny jednostek naukowych (kierownik zarządzający, a następnie ekspertem Ministerstwa Nauki i Szkolnictwa Wyższego). Posiada bogate doświadczenie we współpracy przedsiębiorstw z jednostkami naukowymi oraz transferu, rozwoju i wdrażania innowacyjnych technologii. Pełnił m.in. funkcję Dyrektora Pionu Płatności Elektronicznych w Mennicy Polskiej S.A., a także funkcję Dyrektora ds. Inwestycji dla Grupy Wirtualna Polska. Zasiadał również w szeregu Rad Nadzorczych spółek nowych technologii, w tym podmiotów notowanych na GPW. W VIVE odpowiada m.in. za opracowanie i wdrożenie technologii produkcji kompozytu tekstylnego, wykorzystanie SI w procesach produkcyjnych i wiele innych projektów z obszaru recyklingu.

