

POLITECHNIKA WROCŁAWSKA  
WYDZIAŁ ELEKTRONIKI

---

KIERUNEK: Automatyka i Robotyka (AIR)  
SPECJALNOŚĆ: Robotyka (ARR)

**PRACA DYPLOMOWA  
INŻYNIERSKA**

System wizualnego sortowania  
odpadów plastikowych

a system for visual sorting plastic waste

AUTOR:  
Filip Adamcewicz

PROMOTOR PRACY:  
Dr. Witold Paluszyński,  
Katedra Cybernetyki i Robotyki

# Spis treści

<b>Wstęp</b>	<b>1</b>
<b>1 Przegląd literatury</b>	<b>2</b>
1.1 Stosowane technologie . . . . .	2
1.2 Zbiór danych . . . . .	4
<b>2 Zbiór danych - Plasor</b>	<b>5</b>
2.1 Zdjęcia autora . . . . .	5
2.2 Baza Wadaba . . . . .	6
2.2.1 Ograniczenie bazy . . . . .	6
2.2.2 Odpad o indeksie 0016 . . . . .	6
2.3 Połączenie w Plasor . . . . .	7
2.4 Opisy klas . . . . .	8
2.4.1 PET . . . . .	8
2.4.2 HDPE . . . . .	9
2.4.3 LDPE . . . . .	9
2.4.4 PP . . . . .	10
2.4.5 PS . . . . .	10
2.4.6 Other . . . . .	11
<b>3 Eksperymenty i wyniki</b>	<b>12</b>
3.1 Maszyna wektorów nośnych - SVM . . . . .	12
3.1.1 Implementacja klasyfikatora . . . . .	12
3.1.2 Eksperymenty . . . . .	12
3.1.3 Wnioski . . . . .	16
3.2 Klasyfikator uczenia głębokiego . . . . .	16
3.2.1 Implementacja . . . . .	16
3.2.2 Eksperymenty . . . . .	17
3.2.3 Wnioski . . . . .	20
3.3 Porównanie SVM i CNN . . . . .	20
<b>Podsumowanie</b>	<b>21</b>

# **Wstęp**

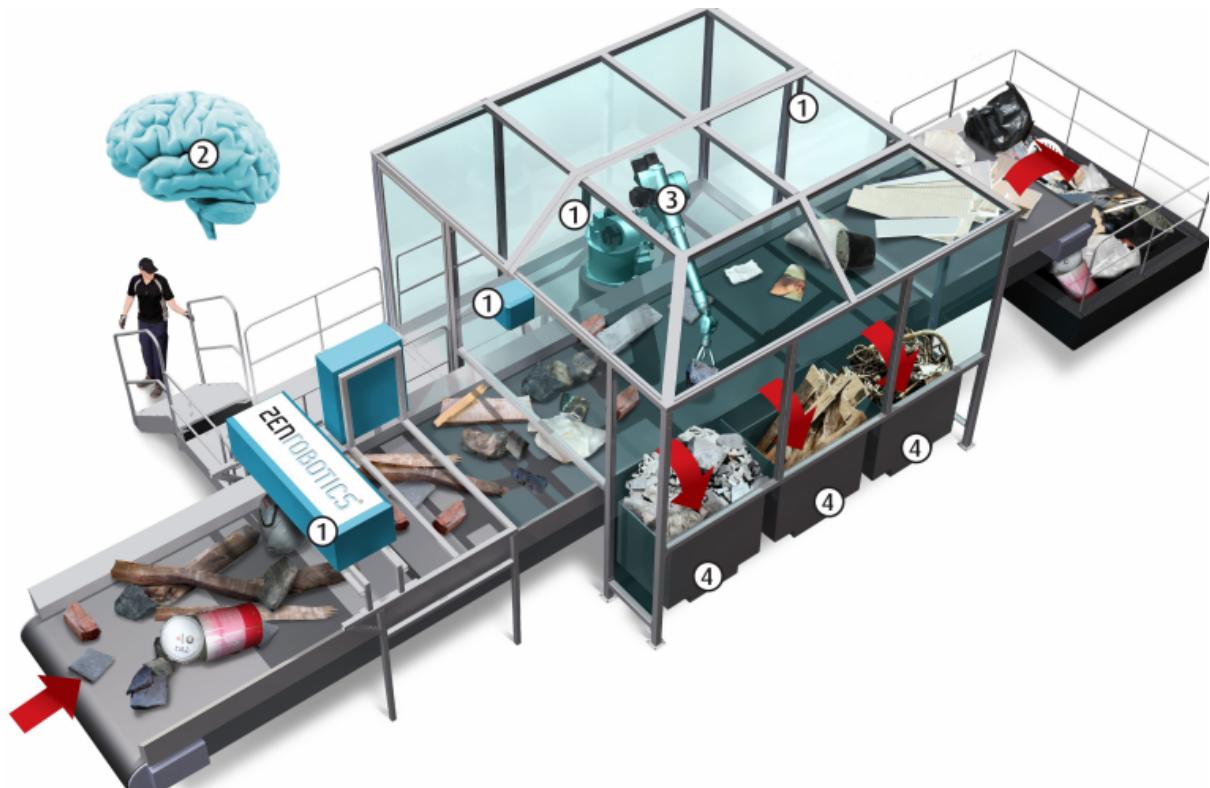
Projekt polega na zastosowaniu uczenia maszynowego do wizualnego rozpoznawania rodzajów odpadów plastikowych. Zbudowany system będzie analizował zdjęcia pojedynczych odpadów i przypisywał im rodzaj odpowiedniego tworzywa sztucznego. Założeniem jest wykorzystanie zbioru danych z internetu oraz rozszerzenie go o własne zdjęcia.

# Rozdział 1

## Przegląd literatury

### 1.1 Stosowane technologie

Wykorzystywane jest wiele metod sortowania odpadów plastikowych. Wśród tych optycznych, popularne jest użycie obrazowania widmowego, analizującego materię w szerszym zakresie promieniowania niż widzialne [2]. W ostatnich latach, środowisko naukowe oraz przemysł dostrzegają potencjał zastosowania sztucznej inteligencji do rozpoznawania odpadów [8, 1]. Na rys. 1.1 przedstawiono komercyjnie dostępny system, łączący sensory, kamery, uczenie maszynowe oraz manipulator, aby zastąpić manualny proces odzysku odpadów budowlanych [5].



Rysunek 1.1: System *ZenRobotics*; 1. Czujniki, 2. *ZenRobotics Brain*, 3. Manipulator, 4. Odzyskane surowce

Próby wykorzystania uczenia maszynowego do klasyfikacji zdjęć odpadów najczęściej

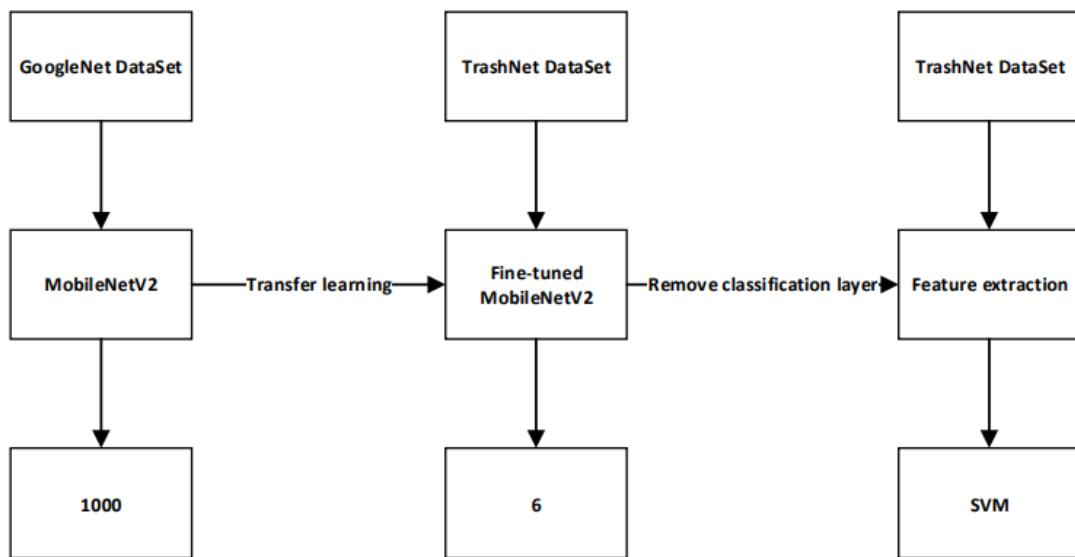
polegają na wykorzystaniu wariantu konwolucyjnych sieci neuronowych [1, 2, 6, 8, 9].

G. Sakr et al. [6] oraz M. Yang, G. Thung [9] porównują klasyfikatory maszyny wektorów nośnych(SVM - ang. *Support Vector Machine*) oraz konwolucyjną sieć neuronową (CNN - ang. *Convolutional Neural Network*), na własnych bazach danych.

Jeden projekt [6] opracował SVM z techniką *bag of features*. Polega ona na wyznaczeniu najczęściej powtarzających się cech obrazów. Następnie opisuje każde zdjęcie jako histogram tych cech. Tak opracowany wektor jest poddawany klasifikacji przez SVM. Drugi zespół [9] również użył tej techniki, jednak do wykrycia cech użył *SIFT* - ang. *Scale-invariant feature transform* na obrazach. Ten proces, korzystając z filtra Laplace'a, pozwolił wydobyć punkty charakterystyczne zdjęć, odporne na skalowanie, szумy oraz oświetlenie.

Odnośnie CNN oba zespoły stworzyły sieć wzorującą się na popularnej strukturze *AlexNet*. Głównymi problemami przy trenowaniu sieci okazał się brak wystarczającej mocy obliczeniowej oraz za mała baza danych zdjęć. SVM osiągnęło dokładność 94.8% [6], 63% [9], a CNN odpowiednio 83% oraz 22%. Mimo to, autorzy wierzą że przy lepszych warunkach CNN osiągnie lepsze wyniki niż tradycyjne metody uczenia maszynowego.

Podobnym projektem porównującym różne metody jest praca Xiujie Xu et al. [8]. Korzystając z udostępnionej bazy *TrashNet* [9], zbadali dokładność siedmiu algorytmów uczenia maszynowego oraz trzech uczenia głębokiego. Najlepszy wynik osiągnął klasyfikator losowych lasów z 64%. Następnie użyta CNN 60% i połączenie analizy głównych składowych z SVM z 57%. Autorzy dokonali jeszcze jednego badania. Wstępnie wytrenowali lekką sieć *MobileNetV2* na bazie *ImageNet*. Następnie zmienili warstwy decyzyjne, dokonali kolejnego treningu, na bazie *TrashNet*, znowu zmienili ostatnie warstwy sieci na SVM i jeszcze raz przetrenowali na *TrashNet*. Proces został zobrazowany na rys. 1.2. w ten sposób udało się uzyskać dokładność 98%.



Rysunek 1.2: Tworzenie modelu sieci zaproponowanego przez Xiujie et al. [8]

J. Bobulski oraz M. Kubanek [2] zaproponowali system do rozpoznawania 4 rodzajów plastiku korzystając z CNN. Po eliminacji innych struktur, ostatecznie porównali dwie: jedną wzorującą się na *AlexNet*, z 23-ma warstwami i drugą, z 15-ma warstwami. Podczas testów zastosowano dwie różne rozdzielcości zdjęć oraz różne stosunki danych treningowych, a testowych. Korzystając z proporcji danych 90% do treningu, 10% testo-

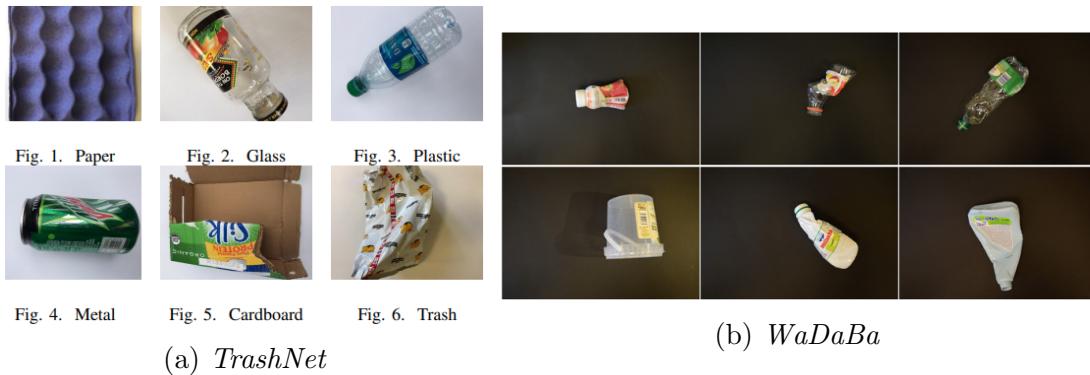
wych udawało się osiągnąć wynik ok. 95% dokładności. Najlepszą konfiguracją z 99.9% dokładności była 15-to warstwowa sieć, przy mniejszym rozmiarze zdjęć.

Zespół C. Bircanoglu et al. [1] zestawił ze sobą różne architektury CNN trenując je na bazie *TrashNet* oraz zaproponował własną strukturę. Najpierw dokonano porównania sieci trenowanych od zera, między innymi *MobileNet*, *InceptionResNetV2* i *DenseNet121*. Na podstawie tego wybrano dwie sieci które zostaną wytrenowane po zastosowaniu na nich metoda *transfer learning* z bazy *ImageNet*. Dzięki temu sieć *DenseNet121* osiągnęła najlepszy wynik ze wszystkich - 95% dokładności. Dodatkowo autorzy zauważają, że klasyfikator działający w przemysłowym systemie musi szybko dokonywać rozpoznawania. Adresując to, opracowali strukturę *RecycleNet*. Jej zaleta polegała na przyśpieszeniu przetwarzania poprzez zmianę wewnętrznych połączeń w sieci *DenseNet121*. Tak wyćwiczona, bez *transfer learningu*, sieć osiągnęła dokładność tylko 81%, lecz była o połowę szybsza w predykcji niż oryginalna wersja.

## 1.2 Zbiór danych

Często powoływana bazą zdjęć odpadów jest *TrashNet*, opracowana oraz udostępniona przez M. Yang, G. Thung [9]. Obejmuje przede wszystkim 5 klas odpadów(papier, metal, szkło, plastik, karton), po ok. 450 na każdy rodzaj. Jest jeszcze 6-ta klasa, *śmieci*, z około stu zdjęć. Przykładowe zdjęcia przedstawiono na rys. 1.3a. Niektórzy badacze zwróciли uwagę na mały rozmiar tej bazy, który trzeba kompensować np. metodami rozszerzania danych(*data augmentation*) [9, 8].

w kontekście samych tworzyw sztucznych bardziej odpowiada baza *WaDaBa* stworzona przez J. Bobulski oraz M. Kubanek [2]. Obejmuje 100 odpadów z 5 rodzajów tworzyw sztucznych. Każdy obiekt jest sfotografowany 40 razy(po 20 zdjęć na jeden sposób oświetlenia), łącznie dając 4000 zdjęć. Przykładowe odpady przedstawiono na rys. 1.3b.



Rysunek 1.3: Przykładowe dane z dwóch baz

w celu kompensacji małej ilości danych, oraz ogólnego poprawienia dokładności rozpoznawania, wykorzystywano *transfer learning*, ćwicząc sieć wstępnie na ogólnej bazie *ImageNet* [8, 1]. Można także rozszerzać dane poprzez zastosowanie transformacji jak rotacje, zmiana kolorów, skalowanie obrazów [1, 2, 6, 8, 9]. w projekcie [2] z niecałych 4000 zdjęć poprzez rotacje uzyskano ok. 120 000. w niektórych przypadkach może to doprowadzić do przeuczenia, jednak ogólnie jest przydatne [1].

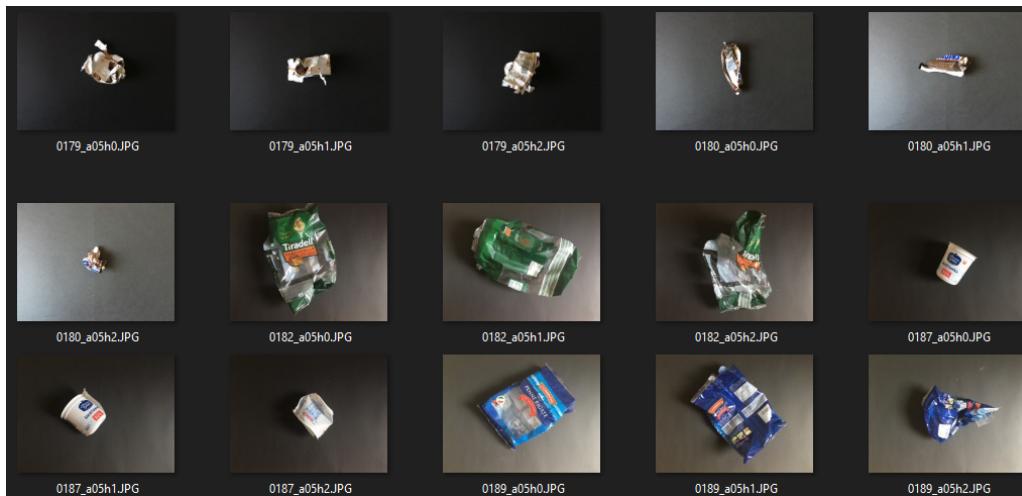
# Rozdział 2

## Zbiór danych - Plasor

Klasyfikatory uczenia maszynowego działają na podstawie danych, przykładów. Im więcej zostanie ich zapewnionych, to algorytm może lepiej się nauczyć przewidywać. Baza Wadaba obejmuje 100 oryginalnych odpadów, co może być niewystarczające dla odpowiedniej generalizacji. Podczas trwania projektu, autor własnoręcznie fotografował domowe odpady z tworzyw sztucznych, aby móc rozszerzyć bazę Wadaba.

### 2.1 Zdjęcia autora

Każdy obiekt został przedstawiony w, średnio, trzech lub czterech różnych stanach. Jest to mniejsza liczba ujęć niż w bazie Wadaba. Różnica wynika z tego, że autor podczas фотографowania subiektywnie decydował kiedy odpad został wystarczająco skatalogowany. Przykładowe zdjęcia wykonane przez autora przedstawiono na rys. 2.1.



Rysunek 2.1: Przykładowe zdjęcia odpadów wykonane przez autora

Zdjęć dokonywano w różnym oświetleniu, najczęściej naturalnym, świecącym z jednej strony. Jako tło użyto dwóch czarnych brystoli. Autor korzystał z tylnego aparatu w telefonie *iPhone 6s*.

Następnie, przy pomocy skryptów napisanych w języku Python, obrazom nadawano nazwę z odpowiednim indeksem obiektu i dwoma parametrami - klasą, zdefiniowaną tak jak w projekcie Wadaba oraz liczbą porządkową ujęcia odpadu.

## 2.2 Baza Wadaba

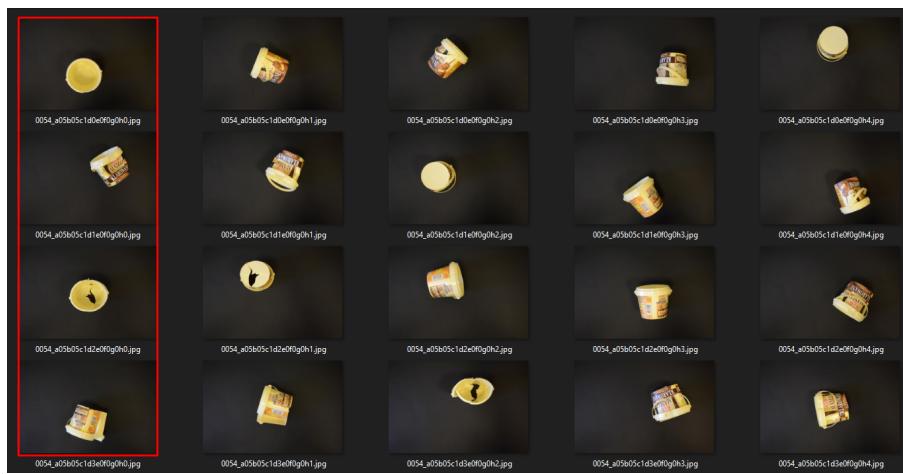
Odpady w bazie Wadaba zostały sfotografowane w 20-stu stanach, przy dwóch różnych sposobach oświetlenia, przykład takiego ujęcia umieszczono na rys. 2.2. w projekcie postanowiono wykorzystać tylko zdjęcia korzystające z oświetlenia przedstawionego po prawej stronie na rysunku.



Rysunek 2.2: Różnica w oświetleniu zdjęcia w bazie Wadaba

### 2.2.1 Ograniczenie bazy

Po próbnych eksperymentach (opisanych w kolejnym rozdziale) i ich analizie, zdecydowano użyć tylko czterech ujęć z każdego obiektu bazy Wadaba. Są to zdjęcia oznaczone parametrem „h0”, które przedstawiają odpad w czterech poziomach deformacji. Na rys. 2.3, pokazano wszystkie 20 obrazów przykładowego obiektu i czerwonym prostokątem zaznaczono 4, które zostały wykorzystane w projekcie. Taka redukcja, zależnie od indywidualnego obiektu, wiąże się z pewną utratą części ujęć, które wnoszą nowe informacje. Jednakże minimalizuje obrazy, które są bardzo podobne do siebie, tak jak np. cztery zdjęcia w prawym dolnym rogu na rys. 2.3. Dodatkowo zapewniony jest podobny stosunek ilości zdjęć na odpad w bazie Wadaba, jaki jest przy zdjęciach zrobionych przez autora.

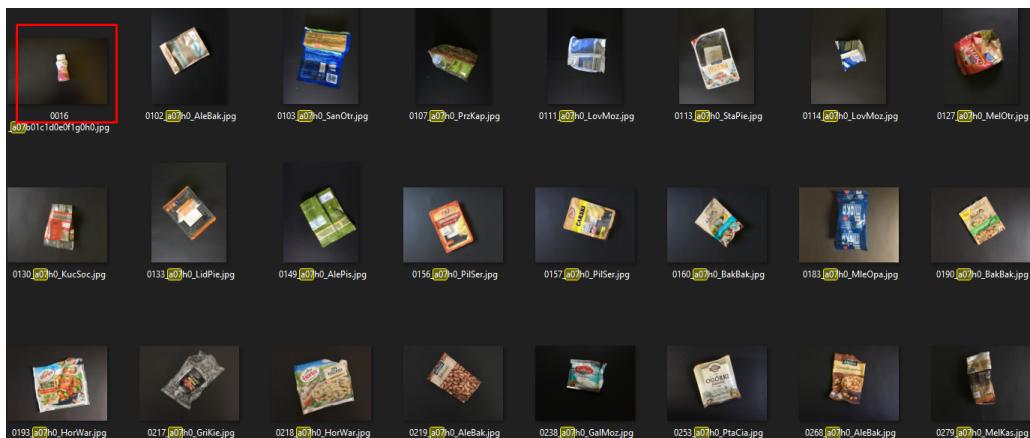


Rysunek 2.3: Przykładowy obiekt z Wadaba, z zaznaczonymi ujęciami o param. „h0”

### 2.2.2 Odpad o indeksie 0016

Dodatkowo autor postanowił zrobić wyjątek i usunąć wszystkie zdjęcia obiektu o indeksie 0016 z bazy Wadaba. Stanowi jedyny z 26-ciu odpadów klasy Other, który przypomina kształtem butelkę. Jest to opakowanie po preparacie odżywczym *Nestle Resource 2.0*

*Fibre*, który, według badania autora, można kupić tylko w aptece. Przez to został uznany za obiekt mocno zaburzający reprezentację tej klasy. Rys. 2.4 przedstawia odpad 0016, zaznaczony czerwonym kwadratem, na tle innych z rodzaju Other.



Rysunek 2.4: Zestawienie obiektu 0016 z innymi z klasy Other

## 2.3 Połączenie w Plasor

Połączono wybrane zdjęcia z bazy Wadaba, wraz z tymi zrobionymi przez autora tworząc bazę Plasor, od ang. *Plastic Sorting Robots*.

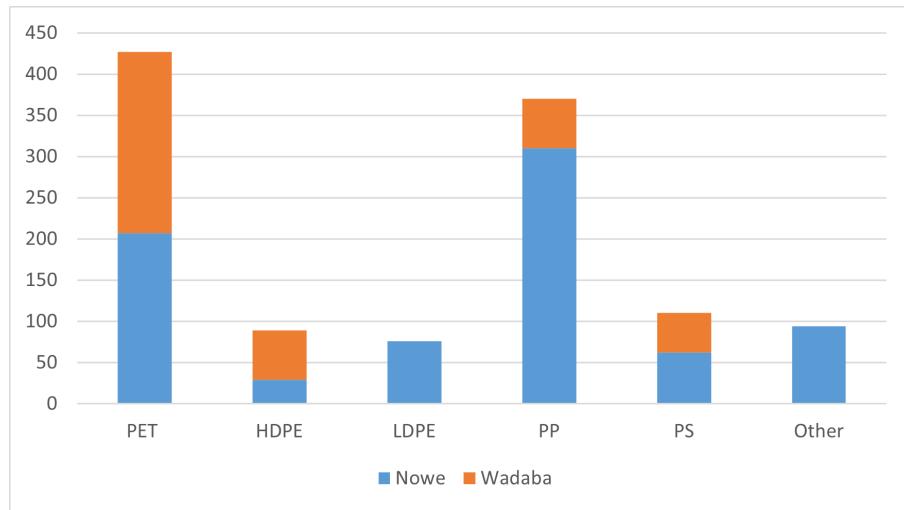
Dalsze prace w ramach projektu będą dokonywane na tym zbiorze. Podsumowując, Plasor składa się z 1166 zdjęć przedstawiających 306 obiektów, w ramach sześciu klas odpadów plastikowych:

- PET
- HDPE
- LDPE
- PP
- PS
- Other

Szczegółową statystykę wspólnej bazy, przedstawiono w tabeli 2.1 oraz na rys. 2.5.

klasa	obiekty	zdjęcia	zdjęcia na obiekt	nowe zdjęcia	zdjęcia Wadaba
PET	111	427	4	207	220
HDPE	27	89	3	29	60
LDPE	19	76	4	76	0
PP	96	370	4	310	60
PS	28	110	4	62	48
Other	25	94	4	94	0
Suma:	306	1166		778	388

Tabela 2.1: Zestawienie informacji o zbiorze Plasor



Rysunek 2.5: Wykres ilości zdjęć na daną klasę

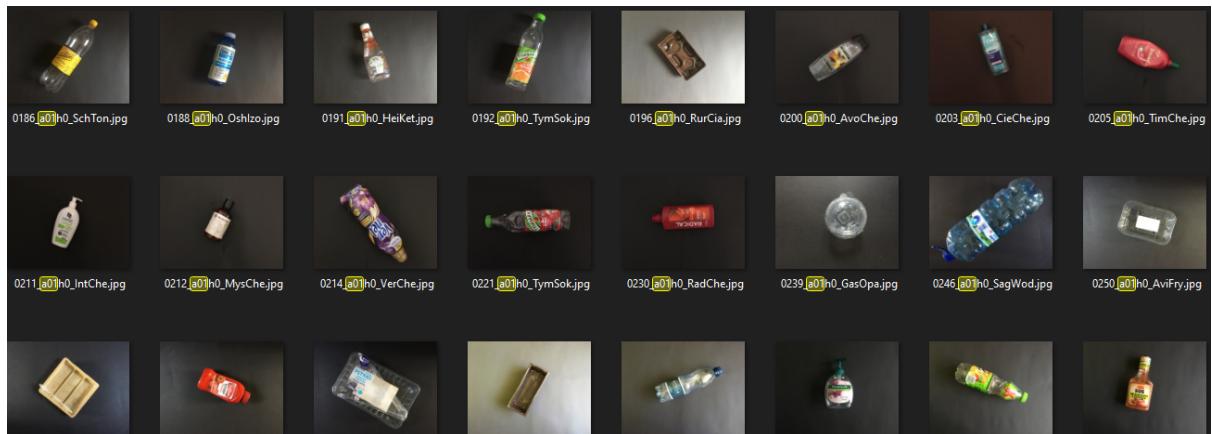
Każdy obraz został zredukowany do rozmiaru 512 na 512 pikseli. Pozwoliło to osiągnąć kompromis pomiędzy wydajnością obliczeniową, a utratą szczegółów fotografii.

## 2.4 Opisy klas

Każdy rodzaj tworzyw sztucznych posiada swoje charakterystyczne cechy. Poniżej znajduje się opis każdej klasy, z obrazami przykładowych obiektów. Ma to pomóc przy późniejszej interpretacji wyników eksperymentów.

### 2.4.1 PET

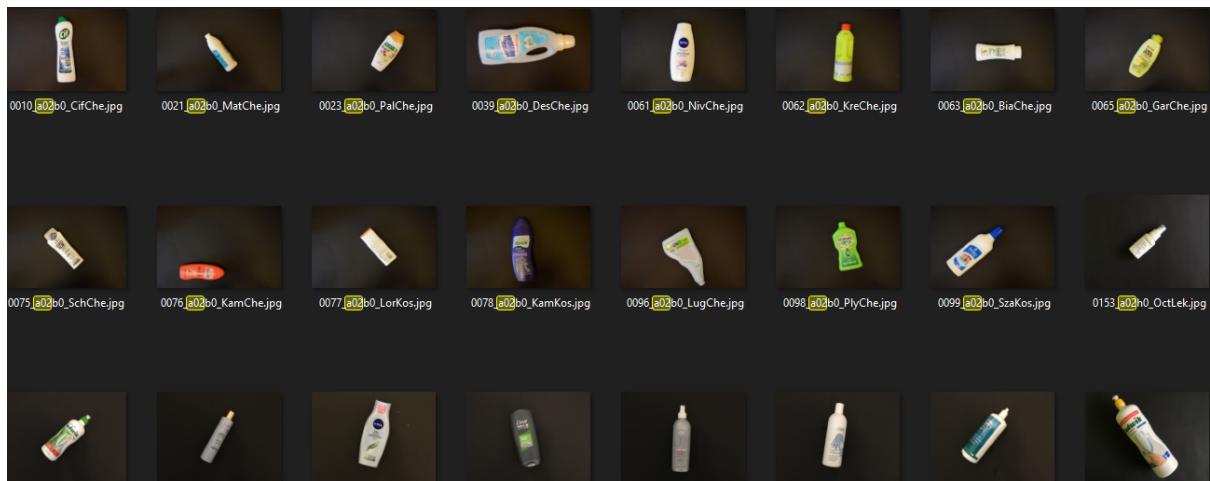
Odpady typu PET składają się w większości z przezroczystych butelek po napojach. Dodatkowo w tej grupie znajdują się też nieprzezroczyste(lub z etykietą) butelki, w tym po chemii jak płynne mydło, czy płyny do płukania prania. Około 10 % zbioru PET w Plasor stanowią obiekty inne niż butelki - np. sztywne „foremkę” po ciastkach, czy jajkach.



Rysunek 2.6: Przykładowe odpady PET

## 2.4.2 HDPE

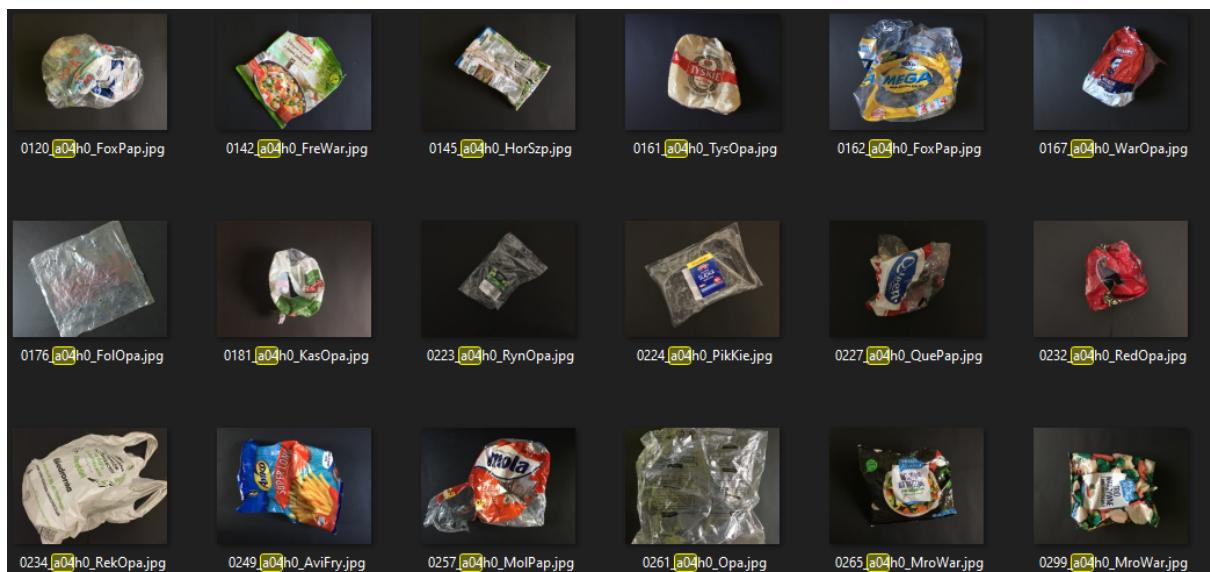
Skatalogowane HDPE to nieprzezroczyste butelki po chemii, kosmetykach. Tylko dwa obiekty mają inne przeznaczenie(płynne lekarstwo, sos do jedzenia), ale ich opakowania też są z grubszego, nieprzezroczystego, plastiku o kształcie butelki.



Rysunek 2.7: Przykładowe odpady HDPE

## 2.4.3 LDPE

LDPE to różnego rodzaju folie, paczki zbiorcze, po np. mrożonych warzywach, sześciopakach, ręcznikach papierowych. Odpady mogą być zarówno przezroczyste lub nie. Nieruszone zajmują dużo powierzchni, najczęściej w postaci prostokąta, jednak po dłuższym czasie bycia zgniecionym odwijają się w różne kształty, szczególnie będąc naderwane.

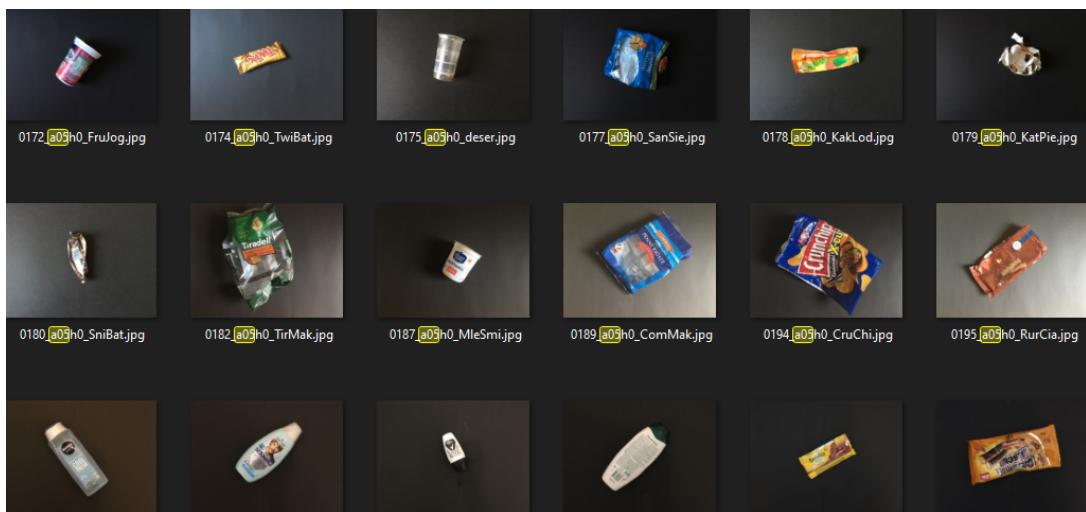


Rysunek 2.8: Przykładowe odpady LDPE

#### 2.4.4 PP

PP jest najbardziej różnorodną klasą. w tej kategorii zebrano m.in. kubeczki po nabiale, opakowania na wynos, jak w PS; worki foliowe podobne do Other oraz LDPE; butelki od kosmetyków, niczym w HDPE, czy PET. Poza tym PP obejmuje też bardziej sztywne opakowania, po mnie, lub foremki po ciastkach.

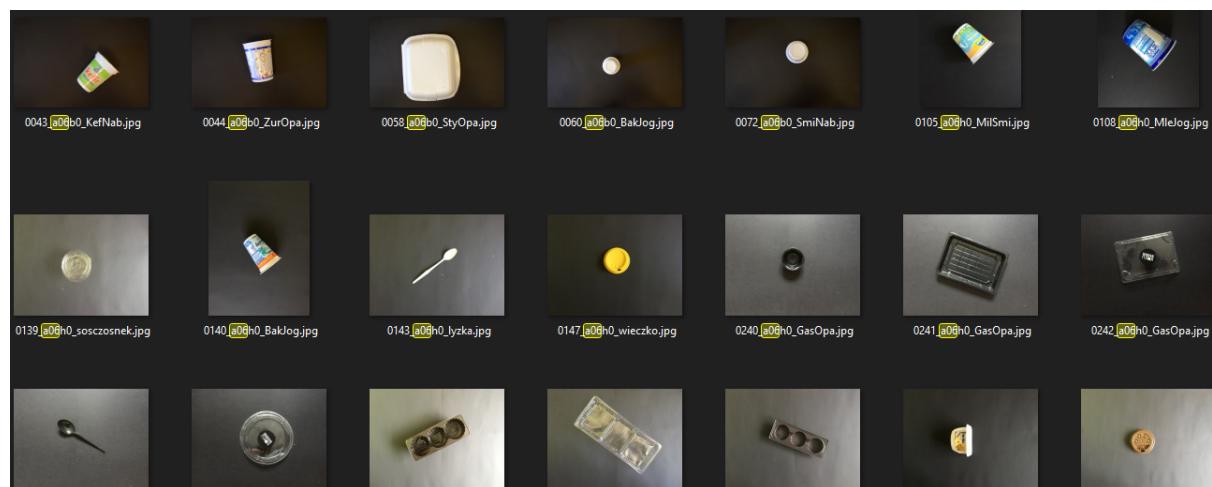
Szczególnie folie występują w kilku rodzajach. Obejmują przezroczyste, bardziej sztywne, do np. makaronu, jak i kolorowe, miękkie oraz małe po batonach, lodaach. Mogą być również nieprzezroczyste, a niektóre nawet błyszczące, „metaliczne” w środku, jak po paczkach chipsów, czy batonach.



Rysunek 2.9: Przykładowe odpady PP

#### 2.4.5 PS

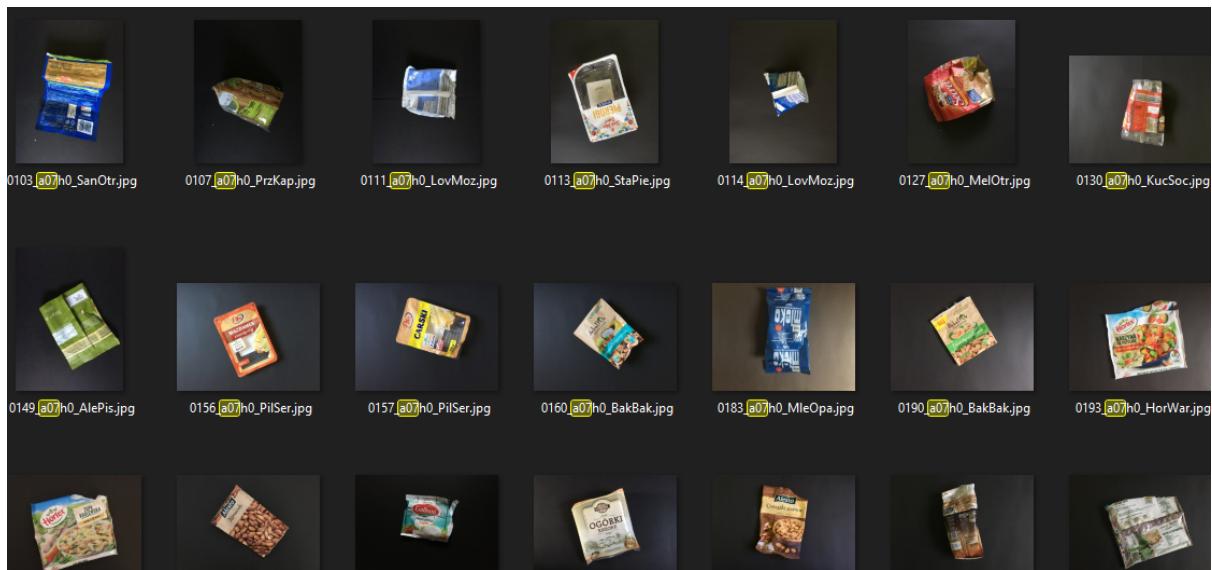
PS które zostały skatalogowane, najczęściej składają się ze sztywnych, kruchych, opakowań. Ich przeznaczeniem jest żywność, głównie nabiał trzymany w kubkach oraz wszelki plastik do jedzenia na wynos - opakowania na posiłki, łyżeczki, wieczka do napojów. Ten rodzaj może być kolorowy, jak przezroczysty.



Rysunek 2.10: Przykładowe odpady PS

### 2.4.6 Other

Other stanowi grupę odpadów z tworzyw sztucznych, które chemiczną kompozycją nie pasują do pozostałych kategorii [7]. Zebrane przez autora obiekty to najczęściej różnorodne nieprzezroczyste opakowania foliowe. Niektóre są sztywne, błyszczące w środku, po orzechach, inne to miękkie folie po mrożonych warzywach, podobnie jak w klasie LDPE. Znajdują się tu też lekko sztywne, przezroczyste, opakowania po jedzeniu z lodówek - serach, gotowych pierogach.



Rysunek 2.11: Przykładowe odpady Other

# Rozdział 3

## Eksperymenty i wyniki

### 3.1 Maszyna wektorów nośnych - SVM

#### 3.1.1 Implementacja klasyfikatora

w pierwszym eksperymencie wzorowano się na pracach [6, 9], w szczególności projekcie Sakr et. al. [6], zaczynając badania od klasyfikatora typu SVM.

Maszyna wektorów nośnych jest algorytmem sprawdzonym, o mocnych podstawach teoretycznych. Potrafi dobrze generalizować na małej ilości próbek (w Plasor niektóre klasy składają się z ok. 25 obiektów) oraz stara się zachować maksymalny margines bezpieczeństwa.

Cały program został napisany w *Matlabie R2017b*. Klasyfikator został wyćwiczony korzystając z funkcji `trainImageCategoryClassifier`. Podstawę implementacji stanowił artykuł z *MathWorks* [3].

#### Warstwowa walidacja krzyżowa

Eksperymenty przeprowadzano używając 10-krotną warstwową walidację krzyżową. Zastosowano podział 90% losowych zdjęć z każdej klasy w zbiorze uczącym i pozostałe 10% w zbiorze testowym.

Podawane wyniki w macierzach błędu są średnią wyników z wszystkich iteracji danego eksperymentu.

#### *Bag of features*

Przed podaniem zdjęć do klasyfikatora, najpierw przetworzono je funkcją `bagOfFeatures`. Funkcja analizuje wszystkie cechy występujące na obrazach uczących. Następnie wybiera najczęstsze 80% i grupuje je w 500 cech algorytmem *K-means*. Na końcu zwraca obiekt pozwalający kodować zdjęcia w wektor o długości tych 500 cech. w takiej postaci obiekty przekazywane są do uczenia, a potem walidacji, klasyfikatora.

#### 3.1.2 Eksperymenty

##### Pierwszy eksperiment

Na samym początku przeprowadzono naukę klasyfikatora na połączonej bazie obrazów autora oraz Wadaba, przed zastosowaniem modyfikacji opisanych w poprzednim rozdziale (zachowując stosunek 20-stu zdjęć na obiekt oraz odpad o indeksie 0016). Wszystkie

funkcje zostały uruchomione z domyślnymi parametrami. Macierz błędów klasyfikacji, zbioru testowego, przedstawiono w tabeli 3.1.

KNOWN	PREDICTED					
	HDPE	LDPE	Other	PET	PP	PS
HDPE	<b>0.13</b>	0.51	0.36	0.00	0.00	0.00
LDPE	0.00	<b>0.72</b>	0.28	0.00	0.00	0.00
Other	0.00	0.51	<b>0.47</b>	0.00	0.00	0.02
PET	0.01	0.58	0.34	<b>0.05</b>	0.00	0.01
PP	0.01	0.59	0.36	0.00	<b>0.01</b>	0.02
PS	0.04	0.54	0.38	0.00	0.00	<b>0.04</b>

\* Average Accuracy is **0.24**

Tabela 3.1: Wynik pierwszego eksperymentu

Najbardziej dominującą klasą jest LDPE, a za nią trochę słabiej Other, pomimo małej liczebności zdjęć obu tych klas. Warto zauważyć, że w bazie Wadaba nie było żadnych obiektów rodzaju LDPE i tylko jeden typu Other.

Na podstawie tego, oraz kilku innych eksperymentów, autor doszedł do modyfikacji bazy Wadaba opisanych w poprzednim rozdziale. Dzięki nim uzyskano lepsze, bardziej wiarygodne, wyniki. Postanowiono uznać te zmiany jako podstawową bazę i nadać jej nazwę Plasor. Wszelkie dalsze eksperymenty operują już niej.

### Podstawowy wynik

Wytrenowano klasyfikator na zbiorze Plasor, zachowując domyślne parametry funkcji. Wyniki znajdują się w tabeli 3.2.

KNOWN	PREDICTED					
	HDPE	LDPE	Other	PET	PP	PS
HDPE	<b>0.83</b>	0.02	0.01	0.03	0.02	0.08
LDPE	0.03	<b>0.79</b>	0.05	0.06	0.05	0.03
Other	0.03	0.29	<b>0.37</b>	0.05	0.08	0.17
PET	0.28	0.13	0.10	<b>0.37</b>	0.01	0.12
PP	0.17	0.32	0.12	0.06	<b>0.16</b>	0.18
PS	0.19	0.01	0.08	0.04	0.03	<b>0.66</b>

\* Average Accuracy is **0.53**

Tabela 3.2: Wynik podstawowego eksperymentu

Klasy LDPE oraz HDPE są dwiema najmniej licznymi, jednak najlepiej rozpoznawalnymi. Dokładność przy tych rodzajach można przypisać temu, że są to najbardziej „jednolite” klasy, w których obiekty mają podobne cechy. Dodatkowo, kiedy SVM dokonuje niepoprawnej klasyfikacji, to najczęściej wybiera te dwa rodzaje. Odpady typu PET są szczególnie mylone jako HDPE. Jest to spójne pod tym względem, że wśród nich znajduje się kilka nieprzezroczystych butelek. Zaś Other oraz PP są mylone z LDPE - warto zauważyć, że te 3 klasy dzielą ze sobą posiadanie opakowań z folii.

Podsumowując, autor wnioskuje, że jednorodność, to jak bardzo obiekty są do siebie podobne w ramach klasy decyduje o dokładności rozpoznawania.

### Parametr *Bag Of Features*

Analizując implementację klasyfikatora, w dokumentacji *MathWorks* [4], znajduje się fragment opisujący funkcję *BagOfFeatures*:

You can extract features based on a feature detector, or you can define a grid to extract feature descriptors. [...] Therefore, use the grid for images that do not contain distinct features, such as an image containing scenery.

Zobrazowanie dwóch sposobów detekcji, przedstawione jest na rys. 3.1. Zielone kwadraty to cechy jakie funkcja wykrywa.



Rysunek 3.1: Porównanie detekcji 'grid', po lewej, oraz 'detector', po prawej.

Wydaje się, że korzystanie z detekcji typu 'grid' może być uprzedzone w kierunku klas z obiektami zajmującymi większą powierzchnię. Przeprowadzono eksperyment, w tych samych warunkach co poprzednio, tylko zmieniając sposób detekcji na 'detector'. Wyniki znajdują się w tabeli 3.3.

PREDICTED

KNOWN	HDPE	LDPE	Other	PET	PP	PS
HDPE	<b>0.65</b>	0.01	0.02	0.09	0.03	0.20
LDPE	0.05	<b>0.52</b>	0.28	0.11	0.03	0.03
Other	0.05	0.30	<b>0.39</b>	0.17	0.03	0.04
PET	0.13	0.13	0.13	<b>0.44</b>	0.05	0.13
PP	0.15	0.19	0.16	0.19	<b>0.09</b>	0.23
PS	0.24	0.05	0.05	0.11	0.10	<b>0.44</b>

\* Average Accuracy is **0.44**

Tabela 3.3: Wyniki po zmianie sposobu detekcji cech

w klasach HDPE, LDPE oraz PS nastąpił spadek dokładności o ok. 20%, a w PP o 8%. Other zmieniło się nieznacznie i tylko PET poprawiło o 7%. Zaprzecza to postawionej tezie, dlatego metoda 'grid' pozostaje jako domyślna opcja w badaniach.

### Próba optymalizacji klas

w różnych eksperymentach najbardziej problematycznymi klasami były PP, Other oraz PS. Cechują się dużą różnorodnością opakowań, podczas ich katalogowania, autorowi zdarzało się błędnie „zgadywać” rozpoznając pomiędzy PP, a PS, lub PP, a Other. Dodatkowo w klasie PET są obiekty, które mają więcej podobieństwa do odpadów PP lub PS, niż butelek PET. W związku z tym, postanowiono połączyć klasy PP, PS, Other, a także, autor subiektywnie wybrał obiekty z PET(17 odpadów), których zdjęcia zostaną połączone w nową klasę: Misc, od ang. *Miscellaneous*. Rodzaj PET został przemianowany na PETb, aby podkreślić że znajdują się tam tylko odpady o kształcie butelek w oryginalnym stanie.

Przy implementacji takiego systemu w zrobotyzowanym stanowisku, odpady klasy Misc byłyby przepuszczane dalej na taśmociągu, po czym można by je poddać specjalnemu traktowaniu.

Sprawdzono skuteczność nowego podziału w eksperymencie z domyślnymi parametrami. Wyniki przedstawiono w tabeli 3.4.

		PREDICTED			
KNOWN		HDPE	LDPE	Misc	PETb
HDPE		<b>0.87</b>	0.08	0.01	0.04
LDPE		0.03	<b>0.81</b>	0.09	0.07
Misc		0.19	0.29	<b>0.40</b>	0.12
PETb		0.20	0.15	0.15	<b>0.51</b>

\* Average Accuracy is **0.65**

Tabela 3.4: Wyniki po przeddefiniowaniu klas

Klasyfikator rozpoznaje klasy HDPE, LDPE o kilka procent lepiej, zaś skuteczność przy klasie PETb wzrosła o 14%. Dokładność przy rodzaju Misc jest podobna do średniej arytmetycznej dokładności z osobnych klas PP, PS, Other. Poprawa może polegać na tym, że klasyfikator unika pomyłek pomiędzy tymi klasami. Jednakże ciągle co trzecie zdjęcie jest mylone z LDPE, a co piąte z HDPE. Także, pomimo zdefiniowania PETb jako „klasy butelek”, są w niej obiekty błędnie klasyfikowane jako LDPE, czy Misc. Podsumowując redefinicja, uproszczenie, klas może trochę pomóc w skutecznym sortowaniu, przynajmniej na początku prac nad projektem.

Zbadano również taki sam klasyfikator, tylko korzystający z detekcji cech 'detector'. Wyniki zawarto w tabeli 3.5.

PREDICTED				
KNOWN	HDPE	LDPE	Misc	PETb
HDPE	<b>0.70</b>	0.00	0.16	0.15
LDPE	0.03	<b>0.58</b>	0.19	0.20
Misc	0.20	0.21	<b>0.34</b>	0.25
PETb	0.11	0.14	0.13	<b>0.63</b>

\* Average Accuracy is **0.56**

Tabela 3.5: Wyniki optymalizacji klas, po zmianie sposobu detekcji cech

Podobnie do eksperymentu na niemodyfikowanym Plasor, rozpoznawalność pogorszyła się we wszystkich klasach, poza PETb, gdzie uległa polepszeniu. Sugeruje to możliwość optymalizacji sortowania, poprzez połączenie klasyfikatora korzystającego z 'grid' dla większości klas, z takim który korzysta z 'detector' dla PETb.

### 3.1.3 Wnioski

Na podstawie wykonanych eksperymentów można podsumować, że SVM najlepiej rozpoznaje klasy LDPE, HDPE z dokładnością ok. 80%; Other, PET, PS ok. 50%; a PP ok. 10%. Widać, że klasyfikator lepiej sobie radzi przy klasach, które są bardziej jednorodne, gdzie wszystkie odpady mają wiele wspólnych cech.

Możliwe, że inne zdefiniowanie klas polepszyłoby skuteczność, jednak raczej nie rozwiązałoby wszystkich problematycznych przypadków.

## 3.2 Klasyfikator uczenia głębokiego

Jako drugi klasyfikator opracowano konwolucyjną sieć neuronową w oparciu o projekty [6] oraz [9]. Jednakże w tych dwóch, oraz innych pracach, częstym problemem dla sieci była mała ilość danych. Dlatego przy implementacji autor wzorował się na artykułach [1], [8], gdzie jest to skompensowane zastosowaniem metody *transfer learning*. Ta metoda sprawia, że sieć neuronowa a priori potrafi rozpoznawać podstawowe kształty. Wystarczy tylko dostosować jej ostatnie warstwy do nowego zadania klasyfikacji.

### 3.2.1 Implementacja

Korzystając z biblioteki *Keras*, zaimplementowano sieć *DenseNet121* wyćwiczoną na zbiorze *ImageNet*, tak jak w opracowaniu [1]. Nie sprecyzowano tam jakie są ostatnie warstwy sieci, dlatego autor samodzielnie je dobrał w serii wstępnych eksperymentów.

Podstawę stanowiła konwolucyjna część sieci *DenseNet121*. Na nią dokładało różne warstwy. Najpierw spróbowano wyuczyć sieć tylko z warstwą wyjściową, lecz CNN szybko przestawała się ulepszać. Podobnie przy dodaniu przed nią warstwy *Dense* o 32 jednostkach. Występował także problem z tym, że dokładność rozpoznawania zbioru uczącego była niższa niż testowego. Ten efekt zmniejszyło dodawanie warstwy *Dropout*. Podczas uczenia sieć wykazywała dużą poprawę rozpoznawalności po pierwszej iteracji, a potem dochodziło do stagnacji. Na to pomogło zmniejszenie współczynnika uczenia optymalizatora.

Ostatecznie użyta sieć po warstwach konwolucyjnych *DenseNet121*, zawiera warstwę *Flatten*, *Dense* z 64-ema jednostkami i aktywacją *ReLU*, a także *Dropout* z parametrem 0.1. Jako warstwę wyjściową zastosowano *Dense* z jednostkami o liczbie klas(6 lub 4) i funkcją aktywacji *softmax*.

Przy komplikacji modelu, funkcja straty liczona jest entropią krzyżową, a także użyto optymalizatora *Adam*, z współczynnikiem szybkości uczenia 0.0001. Rozmiar *batch size* wynosił 32.

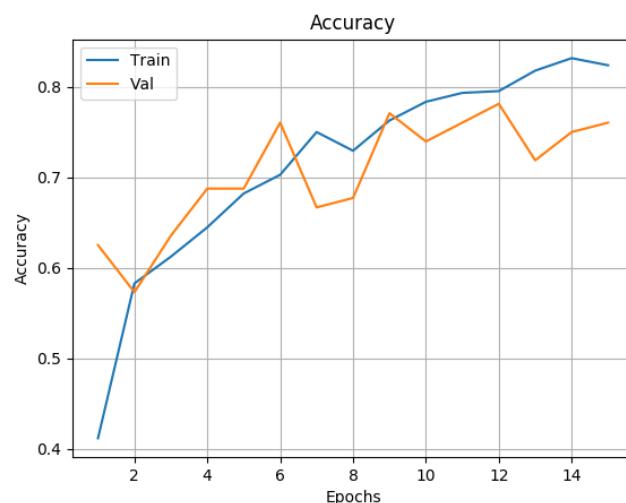
Podczas badań stosowano 10-krotną warstwową walidację krzyżową. Ze względu na wzgłydy praktyczne oraz brak czasu, eksperymenty prowadzono tylko przez 15 epok.

z zakresu rozszerzania danych, rozbudowano proces z [1]. Uzasadnione jest to mniejszym i nierównym zbiorem danych w porównaniu do artykułu. Obrazy uczące były losowo obracane w zasięgu 180-ciu stopni, przerzucane w pionie lub poziomie, a także poddawane transformacji *shear* do 15 stopni. Zdjęcia testowe nie były poddawane obróbce. Operacje zrealizowano korzystając z obiektu klasy `ImageDataGenerator`.

### 3.2.2 Eksperymenty

#### Podstawowy wynik

Najpierw nauczono klasyfikator na zbiorze Plasor, bez modyfikacji. Dokładność walidacji krzyżowej wyniosła 73%. Po każdym etapie walidacji program generował wykres dokładności. Wyniki dla przykładowej fazy przedstawiono na rys. 3.2. Dodatkowo, opracowano macierz błędów klasyfikatora, znajdująca się w tabeli 3.6.



Rysunek 3.2: Przykładowy wykres dokładności sieci, na zbiorze Plasor

Wahania wartości dla zbioru testowego mogą wynikać z małej ilości próbek, większość klas jest reprezentowana przez ok. 10 zdjęć. Prawdopodobnie też niektóre hiperparametry oraz architektura nie są zoptymalizowane.

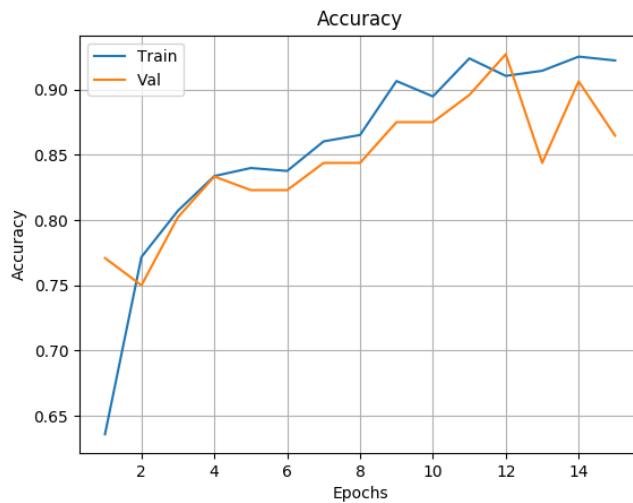
KNOWN	HDPE	LDPE	Other	PET	PP	PS
HDPE	<b>0.78</b>	0.00	0.00	0.22	0.00	0.00
LDPE	0.00	<b>0.88</b>	0.13	0.00	0.00	0.00
Other	0.00	0.22	<b>0.22</b>	0.00	0.56	0.00
PET	0.00	0.00	0.00	<b>0.91</b>	0.05	0.05
PP	0.08	0.03	0.00	0.00	<b>0.78</b>	0.11
PS	0.00	0.00	0.00	0.09	0.45	<b>0.45</b>

Tabela 3.6: Macierz błędu pierwszej sieci

Przy pierwszym teście, na takich samych warunkach, CNN sprawdziła się lepiej niż SVM, w każdej z klas, poza HDPE oraz PS. Jednakże pomylenie obiektów HDPE z, bardziej liczną, klasą PET nie jest zaskakujące. Sieć nauczyła się więcej przykładów, jest uprzedzona ku PET, a te dwie klasy współdzielą nieprzezroczyste butelki ze sobą. Podobne zjawisko zachodzi pomiędzy PS, a liczniejszym PP, które mają wiele podobnych obiektów. Najgorzej wypada klasa Other, mylona z LDPE oraz PP. Ogólnie wyniki można ocenić jako obiecujące.

### Optymalizacja klas

w poprzednim eksperymencie widać było problemy z rozróżnieniem pomiędzy obiektyami Other, PP, PS, a także 10-cioma procentami odpadów PET. Dlatego w ramach drugiego testu wykonano trening sieci na zmodyfikowanym zbiorze Plasor. Powtórzono próbę optymalizacji klas z poprzedniego podziału - połączenie Other, PP, PS i części PET w klasę Misc, a resztę PET przemianowano w PETb. Dokładność poprawiła się do 87%. Wyniki przedstawiono na rys. 3.3 oraz tabeli 3.7.



Rysunek 3.3: Przykładowy wykres dokładności sieci, po optymalizacji klas

Propozycja optymalizacji klas przyczyniła się do ok. 15% poprawy dokładności klasyfikacji. Powstałe wykresy były łagodniejsze w kształcie niż w poprzednim eksperymencie.

KNOWN	HDPE	LDPE	Misc	PETb
HDPE	<b>0.76</b>	0.00	0.16	0.09
LDPE	0.00	<b>0.53</b>	0.48	0.00
Misc	0.03	0.02	<b>0.93</b>	0.03
PETb	0.03	0.00	0.10	<b>0.87</b>

Tabela 3.7: Wyniki sieci po optymalizacji klas

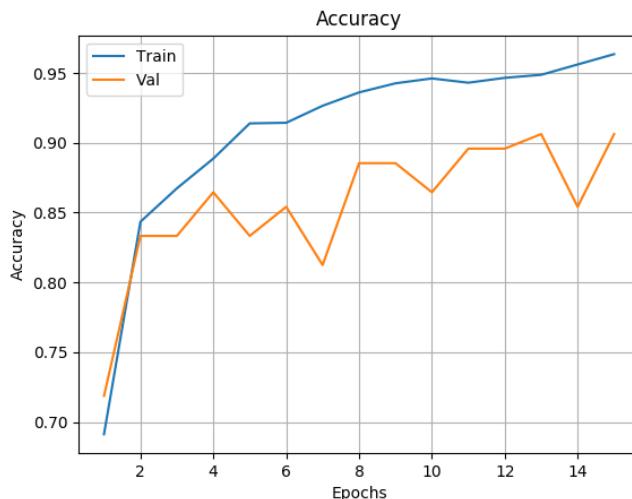
Najbardziej na zmianie ucierpiała rozpoznawalność HDPE oraz LDPE. Warto pamiętać, że z powodu małej liczebności, te klasy są testowane na odpowiednio dziewięciu i osmiu zdjęciach. Dlatego błędna klasyfikacja np. 4 obrazów przekłada się na 50% błędu. PETb pomimo zdefiniowania jako zgrupowanie butelek, ciągle 10% błędnie uznawana jest za Misc. Za to klasa Misc osiągnęła dokładność 93%, wyżej niż którakolwiek ze składowych klas Other, PP, czy PS w poprzednim eksperymencie. Pokazuje to skuteczność takiej fuzji, przynajmniej podczas początkowych etapów projektu.

### Wyrównanie i optymalizacja klas

Ostatni eksperyment dokonano również na Plasor z propozycją optymalizacji, ale też wyrównaną liczebnością klas. Każde zdjęcie ze zbioru uczącego klas HDPE, LDPE oraz PETb zostało powielone odpowiednio 6, 7 i 1 raz. Obrazy w zbiorze testowym pozostały bez zmian. w ten sposób klasa HDPE, LDPE oraz Misc posiadały ok. 560 zdjęć, zaś PETb 656.

z powodu zwiększenia ilości zdjęć tych samych ujęć obiektów, w drodze wyjątku zmieniono parametry modelu klasyfikatora. Współczynnik warstwy regulacji *Dropout* ustalono na 0.2, aby zapobiec zjawisku przeuczenia.

Średnia dokładność wyniosła 90%. Przykładowe wyniki przedstawiono na rys. 3.4 oraz tabeli 3.8.



Rysunek 3.4: Przykładowy wykres dokładności sieci, po optymalizacji i zrównoważeniu Plasor

Pomimo zwiększenia parametru warstwy *Dropout* i tak nasiliło się zjawisko przeuczenia względem poprzednich eksperymentów.

KNOWN	HDPE	LDPE	Misc	PETb
HDPE	<b>0.89</b>	0.00	0.00	0.11
LDPE	0.00	<b>0.75</b>	0.25	0.00
Misc	0.04	0.05	<b>0.89</b>	0.02
PETb	0.05	0.00	0.05	<b>0.90</b>

Tabela 3.8: Wyniki sieci po optymalizacji klas

Powielenie danych pozwoliło uzyskać lepszą dokładność we wszystkich klasach, poza małym pogorszeniem Misc. Te wyniki należałyby jeszcze skonfrontować przy klasyfikacji nowych obiektów. Jednak metoda powielenia i transformacji zdjęć, zdecydowanie pomaga przy mniej licznych klasach.

### 3.2.3 Wnioski

Pomimo relatywnie małej ilości danych klasyfikator CNN uzyskał dobre wyniki. Można to zawdzięczać metodzie *transfer learning* - autor dodatkowo sprawdził przykładową, prostą architekturę z internetu i nie poradziła sobie w połowie tak dobrze jak te opisane. Ważnym aspektem sieci okazało się zadbanie o równowagę reprezentacji każdej klasy. Ostatecznie osiągnięto dokładność w okolicach 80-90% dla każdej klasy.

Poza zbiorem danych, w budowie samego modelu sieci jest wiele hiperparametrów, które w różny sposób wpływają na klasyfikator. Dobranie odpowiednich wartości może wiązać się z wieloma, kosztownymi obliczeniowo, eksperymentami.

## 3.3 Porównanie SVM i CNN

Silnymi stronami klasyfikatora SVM jest możliwość operowania na małej, niewyrównanej, ilości danych oraz rozpoznawanie klas złożonych z podobnych obiektów. Sieci neuronowe są bardziej wymagające, pod względem danych, swojej budowy, mocy obliczeniowej. Jednak dzięki nim udało się uzyskać bardziej obiecujące wyniki.

w obu algorytmach można zauważać problematyczne przypadki odpadów. Poza rozszerzeniem zbioru danych, autor widzi trzy sposoby na polepszenie dokładności:

1. Zmiana definicji klas, stworzenie grup jak np. „kosmetyki”, „folie”.
2. Zastosowanie fuzji czujników. Połączenie obrazu z kamer z analizą spektralną, czy sensorami taktylnymi.
3. Połączenie różnych klasyfikatorów wyspecjalizowanych w innych klasach.

# Podsumowanie

Udało się zrealizować wszystkie postawione zadania. Znaleziono zbiór danych w internecie i rozbudowano go o dodatkowe obiekty, tworząc nową bazę odpadów z tworzyw sztucznych.

Sortowanie odpadów z tworzyw sztucznych jest dość trudnym zagadnieniem. Obiekty potrafią zarówno być różnorodne w ramach jednego rodzaju, jak i występują podobieństwa pomiędzy klasami. Każdy algorytm, z różnymi parametrami, radzi sobie lepiej lub gorzej, w danych warunkach.

Ostatecznie opracowano dwa klasyfikatory uczenia maszynowego potrafiące rozpoznawać zdjęcia pojedynczych odpadów:

1. Maszyna wektorów nośnych z techniką *bag of features* - sprawdzony algorytm, który uzyskał średnio 60% dokładności. w szczególności skuteczny przy prostych klasach.
2. Konwolucyjna sieć neuronowa, z wykorzystaniem metody *transfer learning* - nowoczesne podejście, które eksperymentalnie dochodzi do 90% dokładności. Wymagający, ale bardziej obiecujący klasyfikator.

# Bibliography

- [1] Cenk Bircanoglu et al. “Recyclenet: Intelligent waste sorting using deep neural networks”. In: *2018 Innovations in Intelligent Systems and Applications (INISTA)*. IEEE. 2018, pp. 1–7.
- [2] Janusz Bobulski and Mariusz Kubanek. “Waste classification system using image processing and convolutional neural networks”. In: *International Work-Conference on Artificial Neural Networks*. Springer. 2019, pp. 350–361.
- [3] *Image Category Classification Using Bag of Features*. 2020 (accessed 30.11.2020). URL: <https://www.mathworks.com/help/vision/ug/image-category-classification-using-bag-of-features.html>.
- [4] *Image Classification with Bag of Visual Words*. 2020 (accessed 03.12.2020). URL: <https://www.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html>.
- [5] Tuomas J Lukka et al. “ZenRobotics Recycler—Robotic sorting using machine learning”. In: *Proceedings of the International Conference on Sensor-Based Sorting (SBS)*. 2014, pp. 1–8.
- [6] George E Sakr et al. “Comparing deep learning and support vector machines for autonomous waste sorting”. In: *2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET)*. IEEE. 2016, pp. 207–212.
- [7] *Wikipedia Recycling codes*. 2020 (accessed 07.12.2020). URL: [https://en.wikipedia.org/wiki/Recycling\\_codes](https://en.wikipedia.org/wiki/Recycling_codes).
- [8] Xiujie Xu, Xuehai Qi, and Xingjian Diao. “Reach on Waste Classification and Identification by Transfer Learning and Lightweight Neural Network”. In: (2020).
- [9] Mindy Yang and Gary Thung. “Classification of trash for recyclability status”. In: *CS229 Project Report* 2016 (2016).