



Politechnika  
Śląska

## PROJEKT INŻYNIERSKI

System wizyjny dla robota mobilnego

Łukasz GRABARSKI

Nr albumu: 300434

**Kierunek:** Automatyka i Robotyka

**Specjalność:** Technologie Informacyjne

**PROWADZĄCY PRACĘ**

dr inż. Krzysztof Jaskot

**KATEDRA** Automatyki i Robotyki

**Wydział** Automatyki, Elektroniki i Informatyki

Gliwice 2025



**Tytuł pracy**

System wizyjny dla robota mobilnego

**Streszczenie**

(Streszczenie pracy – odpowiednie pole w systemie APD powinno zawierać kopię tego streszczenia.)

**Słowa kluczowe**

RaspberryPi, Python, OpenCV, YOLO, PyTorch

**Thesis title**

Vision system for a mobile robot

**Abstract**

(Thesis abstract – to be copied into an appropriate field during an electronic submission – in English.)

**Key words**

RaspberryPi, Python, OpenCV, YOLO, PyTorch



# Spis treści

<b>1 Wstęp</b>	<b>1</b>
1.1 Wprowadzenie w problem . . . . .	1
1.2 Osadzenie problemu w dziedzinie . . . . .	2
1.3 Cel pracy . . . . .	2
1.4 Zakres pracy . . . . .	3
1.5 Struktura pracy . . . . .	3
1.6 Wkład własny autora . . . . .	4
<b>2 Analiza tematu i przegląd literatury</b>	<b>5</b>
2.1 Sformułowanie problemu . . . . .	5
2.2 Stan wiedzy i osadzenie w kontekście aktualnych badań . . . . .	6
2.2.1 Metody przetwarzania obrazu . . . . .	6
2.2.2 Algorytmy detekcji obiektów . . . . .	6
2.2.3 Algorytm YOLO . . . . .	6
2.2.4 Zastosowania w robotyce mobilnej . . . . .	7
2.3 Studia literaturowe i znane rozwiązania . . . . .	7
2.3.1 System programowania i sterowania robota mobilnego . . . . .	8
2.3.2 Aktywny system wizyjny dla robota kroczącego . . . . .	8
2.3.3 System wspomagania nawigacji osób niewidomych . . . . .	9
<b>3 Przygotowanie i trening modelu detekcji obiektów</b>	<b>11</b>
3.1 Przygotowanie danych wzorcowych . . . . .	11
3.2 Platforma Roboflow . . . . .	12
3.2.1 Proces użycia witryny Roboflow . . . . .	13
3.2.2 Trening zestawu danych . . . . .	15
3.3 Trening modelu YOLOv7 . . . . .	17
3.3.1 Przygotowanie środowiska . . . . .	17
3.3.2 Instalacja wymaganych bibliotek . . . . .	17
3.3.3 Pobranie i przygotowanie danych treningowych . . . . .	17
3.3.4 Przygotowanie modelu . . . . .	18
3.3.5 Trening modelu . . . . .	18

<b>4 Proces tworzenia robota mobilnego</b>	<b>19</b>
4.1 Założenia logiczne . . . . .	19
4.2 Konstrukcja robota mobilnego . . . . .	21
4.2.1 Wykorzystane moduły . . . . .	21
4.2.2 Schemat elektryczny - dokonczyc . . . . .	23
4.2.3 Projekt robota mobilnego . . . . .	23
4.3 Dostosowanie RaspberryPi . . . . .	26
4.3.1 Spis wymaganych bibliotek . . . . .	27
<b>5 Analiza działania systemu wizyjnego</b>	<b>29</b>
5.1 Detekcja obiektu za pomocą narzędzi głębokiego uczenia . . . . .	29
5.1.1 Opis działania programu . . . . .	29
5.1.2 Uzyskane rezultaty . . . . .	30
5.2 Detekcja obiektu za pomocą binaryzacji i klasycznych technik przetwarzania obrazów . . . . .	30
5.3 Porównanie skuteczności algorytmów . . . . .	30
<b>6 Problematyka i rozwiązania</b>	<b>31</b>
<b>7 Podsumowanie i wnioski</b>	<b>33</b>
<b>Bibliografia</b>	<b>35</b>
<b>Spis skrótów i symboli</b>	<b>39</b>
<b>Źródła</b>	<b>41</b>
<b>Lista dodatkowych plików, uzupełniających tekst pracy</b>	<b>43</b>
<b>Spis rysunków</b>	<b>45</b>
<b>Spis tabel</b>	<b>47</b>

# Rozdział 1

## Wstęp

Rozwój technologii komputerowej oraz metod sztucznej inteligencji w ostatnich dekadach znacząco zmienił podejście do projektowania i implementacji systemów wizyjnych. Szczególnie istotne w tym kontekście stały się algorytmy głębokiego uczenia, które umożliwiły precyzyjne rozpoznawanie i analizę obrazów w czasie rzeczywistym. W połączeniu z dostępnością wydajnych i ekonomicznych platform obliczeniowych, takich jak Raspberry Pi, rozwiązania te znajdują szerokie zastosowanie w robotyce mobilnej. Niniejsza praca koncentruje się na opracowaniu systemu wizyjnego dla robota mobilnego z wykorzystaniem algorytmów YOLOv7, biblioteki PyTorch oraz OpenCV. System ma umożliwić detekcję i śledzenie obiektów w czasie rzeczywistym, co stanowi kluczowy element autonomicznego działania robota w dynamicznym środowisku.

### 1.1 Wprowadzenie w problem

Robotyka mobilna jest jedną z najszybciej rozwijających się dziedzin technologii, znajdującą zastosowanie zarówno w przemyśle, jak i w gospodarstwach domowych. Dzięki zastosowaniu systemów wizyjnych, roboty mobilne zyskują zdolność interakcji z otoczeniem, co zwiększa zakres ich funkcjonalności. Systemy wizyjne umożliwiają identyfikację obiektów, analizę ich ruchu, a także podejmowanie decyzji w czasie rzeczywistym, co otwiera drogę do autonomicznego działania robotów w dynamicznych środowiskach.

Rozwój technologii takich jak głębokie uczenie (ang. *Deep Learning*) oraz dostępność platform sprzętowych o dużej mocy obliczeniowej, poczynając od Raspberry Pi aż po NVIDIA Jetson, pozwalają na implementację złożonych algorytmów przetwarzania obrazu i sterowania robotem. W niniejszej pracy wykorzystano otwartoźródłową bibliotekę programistyczną PyTorch oraz algorytm YOLO (*You Only Look Once*) do realizacji zadań związanych z detekcją i śledzeniem obiektów.

## 1.2 Osadzenie problemu w dziedzinie

Identyfikacja i śledzenie obiektów w czasie rzeczywistym jest jednym z kluczowych wyzwań w robotyce mobilnej. Dzięki zastosowaniu metod przetwarzania obrazu, roboty mogą nie tylko „widzieć” otoczenie, ale również analizować je i realizować przypisane im funkcje. Problem ten jest szczególnie istotny w systemach autonomicznych, które muszą działać w zróżnicowanym środowisku i dostosowywać się do dynamicznych i nieprzewidywalnych zmian.

Wyzwaniem w realizacji takich systemów jest optymalizacja algorytmów wizyjnych pod kątem wydajności i dokładności, dostosowując je do ograniczeń sprzętowych platform takich jak Raspberry Pi. Integracja z systemami sterowania i nawigacji robotów mobilnych wymaga również opracowania odpowiednich rozwiązań programistycznych oraz architektonicznych.

## 1.3 Cel pracy

Celem pracy jest opracowanie systemu wizyjnego dla robota mobilnego, który umożliwi identyfikację różnego rodzaju obiektów, śledzenie ich ruchu oraz podążanie za nimi. W szczególności praca koncentruje się na:

- Implementacji algorytmu YOLOv7 w oparciu o bibliotekę PyTorch na platformie Raspberry Pi 4B.
- Optymalizacji systemu wizyjnego do pracy w czasie rzeczywistym w różnych warunkach środowiskowych.
- Integracji systemu wizyjnego z układem sterowania robota mobilnego.
- Analizie wydajności systemu oraz ocenie jego funkcjonalności w warunkach rzeczywistych.

## 1.4 Zakres pracy

Realizacja projektu obejmuje następujące etapy:

- **Analiza literatury** – przegląd istniejących rozwiązań w zakresie detekcji i śledzenia obiektów oraz systemów wizyjnych.
- **Projekt systemu** – zaprojektowanie architektury systemu wizyjnego, uwzględniając specyfikę robota mobilnego oraz ograniczenia sprzętowe platformy Raspberry Pi.
- **Implementacja** – realizacja algorytmów detekcji i śledzenia obiektów z wykorzystaniem YOLOv7, PyTorch oraz OpenCV.
- **Integracja** – połączenie systemu wizyjnego z modułem sterowania robota.
- **Testy i walidacja** – przeprowadzenie testów w rzeczywistych warunkach operacyjnych oraz analiza wyników działania systemu.

## 1.5 Struktura pracy

Praca została podzielona na następujące rozdziały:

- Rozdział pierwszy – **Wstęp**: zawiera wprowadzenie w problem, cel pracy, zakres oraz strukturę dokumentu.
- Rozdział drugi – **Analiza tematu**: przedstawia aktualny stan wiedzy, analizę literatury oraz istniejących rozwiązań w dziedzinie systemów wizyjnych.
- Rozdział trzeci – **Wymagania i narzędzia**: opisuje wymagania funkcjonalne systemu oraz narzędzia i technologie użyte w pracy.
- Rozdział czwarty – **Projekt systemu wizyjnego**: przedstawia architekturę systemu, zastosowane algorytmy oraz szczegóły implementacji.
- Rozdział piąty – **Weryfikacja i walidacja**: opisuje metodologię testowania, uzyskane wyniki oraz ich analizę.
- Rozdział szósty – **Podsumowanie i wnioski**: zawiera podsumowanie wykonanej pracy oraz propozycje dalszych badań.

## 1.6 Wkład własny autora

W ramach niniejszej pracy autor:

- Przygotował zestaw danych treningowych zawierający obrazy obiektu w różnych warunkach oświetleniowych.
- Przeprowadził trening sieci neuronowej YOLOv7 na platformie Google Colab.
- Samodzielnie zaimplementował system wizyjny z wykorzystaniem YOLOv7.
- Zaprojektował robota mobilnego oraz moduł sterowania.
- Zintegrował system wizyjny z robotem mobilnym.
- Przeprowadził testy systemu w różnych warunkach środowiskowych oraz dokonał analizy wyników.

# Rozdział 2

## Analiza tematu i przegląd literatury

### 2.1 Sformułowanie problemu

Współczesna robotyka mobilna stawia liczne wyzwania związane z autonomicznym działaniem robotów w dynamicznych środowiskach. Jednym z kluczowych aspektów jest detekcja i śledzenie obiektów w czasie rzeczywistym, co umożliwia robotowi interakcję z otoczeniem oraz podejmowanie decyzji w oparciu o dane wizualne. Problem ten komplikuje się w warunkach zmennego oświetlenia, ograniczonej mocy obliczeniowej oraz konieczności integracji algorytmów wizji komputerowej z systemami sterowania.

Systemy wizyjne pełnią istotną rolę w realizacji autonomii robota, umożliwiając:

- wykrywanie przeszkód i nawigację w środowisku,
- identyfikację i śledzenie określonych obiektów,
- analizę otoczenia w czasie rzeczywistym.

W niniejszej pracy problematyka ta jest analizowana w kontekście systemów wizyjnych opartych na algorytmach głębokiego uczenia, takich jak YOLOv7, zaimplementowanych na platformie Raspberry Pi.

## 2.2 Stan wiedzy i osadzenie w kontekście aktualnych badań

W ostatnich latach w dziedzinie robotyki mobilnej i systemów wizyjnych zaobserwowano dynamiczny rozwój technik opartych na głębokim uczeniu. Metody te, w szczególności algorytmy detekcji obiektów, znalazły zastosowanie w projektach przemysłowych i badaniach naukowych.

### 2.2.1 Metody przetwarzania obrazu

Klasyczne podejścia, takie jak filtry Sobela czy segmentacja obrazów, miały ograniczoną skuteczność w dynamicznych środowiskach. Wprowadzenie algorytmów głębokiego uczenia, takich jak YOLO (*You Only Look Once*), znaczaco poprawiło możliwości detekcji obiektów w czasie rzeczywistym.

### 2.2.2 Algorytmy detekcji obiektów

Współczesne metody detekcji obiektów można podzielić na:

- **Algorytmy jednoetapowe (ang. single-stage)**: YOLO, SSD, które łączą detekcję i klasyfikację w jednym przebiegu, oferując wysoką wydajność w czasie rzeczywistym.
- **Algorytmy dwuetapowe (ang. two-stage)**: Faster R-CNN, które cechuje wyższą dokładność, ale kosztem wydajności.

YOLO wyróżnia się szybkością działania oraz zdolnością do pracy na urządzeniach z ograniczoną mocą obliczeniową, takich jak Raspberry Pi.

### 2.2.3 Algorytm YOLO

Algorytm YOLO (*You Only Look Once*) został zaprojektowany jako zintegrowany system detekcji obiektów w czasie rzeczywistym, przekształcając problem detekcji w zadanie regresji. YOLO stosuje jedną sieć neuronową do analizy całego obrazu i równoczesnego przewidywania współrzędnych ramek ograniczających oraz klas obiektów [redmon2016yolo].

**Założenia projektowe** Model YOLO dzieli obraz na siatkę  $S \times S$ , gdzie każda komórka siatki przewiduje:

- $B$  ramek ograniczających (*bounding boxes*) wraz z ich współrzędnymi  $(x, y, w, h)$ ,
- prawdopodobieństwo obecności obiektu w każdej ramce,

- prawdopodobieństwo klasowe dla każdego obiektu.

Łączne przewidywania są zapisywane jako tensor o wymiarze  $S \times S \times (B \cdot 5 + C)$ , gdzie  $C$  to liczba klas [4].

**Architektura sieci YOLO** Sieć YOLO opiera się na warstwach konwolucyjnych, w których kluczowe znaczenie ma redukcja wymiarów za pomocą warstw  $1 \times 1$  i  $3 \times 3$ . W wersji YOLOv7 wprowadzono szereg usprawnień, takich jak:

- użycie modułów re-parametryzowanych (*re-parameterized convolution*) dla poprawy propagacji gradientu,
- lepsza agregacja cech w warstwach za pomocą strategii E-ELAN (*Extended Efficient Layer Aggregation Networks*),
- optymalizacja procesu uczenia poprzez dynamiczne przypisywanie etykiet (*coarse-to-fine label assignment*) [5].

W ramach tej pracy YOLO zostało wykorzystane do detekcji obiektów w systemie wizyjnym robota mobilnego. Model YOLOv7, dzięki swojej optymalnej architekturze i wydajności, umożliwił realizację systemu wizyjnego z zachowaniem wymagań ograniczonej mocy obliczeniowej platformy Raspberry Pi.

#### 2.2.4 Zastosowania w robotyce mobilnej

Systemy wizyjne oparte na YOLO są wykorzystywane w projektach przemysłowych i naukowych. Przykłady zastosowań obejmują:

- nawigację autonomiczną w robotach AGV (Automated Guided Vehicles),
- wykrywanie przeszkód i analiza ścieżki w pojazdach autonomicznych,
- systemy inspekcji w przemyśle.

### 2.3 Studia literaturowe i znane rozwiązania

W literaturze naukowej można znaleźć liczne prace dotyczące implementacji systemów wizyjnych w robotyce mobilnej. W niniejszym rozdziale przytoczono przykłady wybranych badań.

### 2.3.1 System programowania i sterowania robota mobilnego

W pracy [2] prof. dr hab. inż. Krzysztof Kozłowski i współautorzy przedstawili konцепcję oraz realizację modułowego robota mobilnego, zaprojektowanego do zastosowań transportowych w zamkniętych środowiskach. Celem projektu było opracowanie systemu programowania i sterowania, uwzględniającego wymogi bezpieczeństwa oraz przyjaznego interfejsu użytkownika.

**Cel systemu:** Robot został zaprojektowany jako urządzenie transportowe o nośności do 120 kg, dedykowane do przewozu ładunków w pomieszczeniach zamkniętych. Wymagania projektowe obejmowały:

- Modułową budowę, umożliwiającą integrację różnych podsystemów.
- Zapewnienie bezpieczeństwa użytkowania dzięki systemowi zderzaków i czujników.
- Możliwość programowania trasy przejazdu za pomocą przyjaznego interfejsu.

**Metodologia:** Autorzy zastosowali klasyczny model pojazdu dwukołowego z napędem różnicowym oraz pasywnym kołem podporowym. System sterowania oparto na komputerze pokładowym klasy PC oraz dedykowanym oprogramowaniu zbudowanym w architekturze klient-serwer. Wprowadzono język programowania LeoOS Programming Language (LPL), który umożliwia definiowanie trajektorii ruchu robota za pomocą punktów referencyjnych.

**Wyniki:** Robot został wyposażony w zaawansowane systemy sensoryczne, w tym skaner laserowy oraz sieć sensorów ultradźwiękowych i podczerwieni. Wyniki eksperymentalne wskazują na wysoką precyzję sterowania oraz elastyczność w realizacji zadań, takich jak:

- Dokowanie do źródła zasilania.
- Transport ładunków przy pomocy zautomatyzowanego zaczepu.
- Nawigacja w otoczeniu z przeszkodami.

**Wnioski:** Praca autorów pokazuje potencjał zastosowania modularnych systemów sterowania w robotyce mobilnej, szczególnie w środowiskach o dużych wymaganiach bezpieczeństwa. Implementacja języka LPL oraz architektury modułowej zwiększa elastyczność i efektywność systemu, jednocześnie otwierając drogę do dalszego rozwoju.

### 2.3.2 Aktywny system wizyjny dla robota kroczącego

W pracy [1] mgr inż. Przemysław Łabęcki i Dr hab. inż. Andrzej Kasiński przedstawili aktywny system wizyjny, zaprojektowany z myślą o implementacji w robotach kroczących. System ten opierał się na połączeniu kamery oraz oświetlacza laserowego, którego

zadaniem było rzutowanie światła na otoczenie w celu pozyskania informacji o geometrii przestrzeni.

**Cel systemu:** Głównym celem opisanego systemu było zapewnienie robotowi informacji 3D o przestrzeni znajdującej się bezpośrednio przed nim, co umożliwiało zarówno nawigację, jak i detekcję przeszkód.

**Metodologia:** - Zastosowano arkusz świetlny emitowany przez laser, a jego ślad rejestrowany był przez kamerę. - Opracowano matematyczny model układu, definiujący relacje geometryczne między kamerą, oświetlaczem i otoczeniem. - Przeprowadzono kalibrację systemu, aby dokładnie określić parametry układu, takie jak kąty widzenia kamery i położenie głównych punktów odniesienia. - Do detekcji śladu wiązki lasera na obrazie wykorzystano techniki przetwarzania obrazu w przestrzeni barw HSV.

**Wyniki:** - System pozwalał na odwzorowanie geometrii otoczenia z błędem względnym nieprzekraczającym 5%. - Przeprowadzono eksperymenty z obiektyami znajdującymi się w różnych odległościach od kamery, wykazując wysoką dokładność w pomiarach wymiarów geometrycznych. - System sprawdził się jako dalmierz i narzędzie do tworzenia mapy przestrzeni 3D, co było szczególnie przydatne w planowaniu ruchu robota kroczącego.

**Wnioski:** Wyniki badań potwierdzają przydatność aktywnych systemów wizyjnych w robotyce mobilnej, szczególnie w warunkach wymagających precyzyjnych pomiarów odległości i analizy przestrzeni. Rozwiążanie to cechuje niski koszt i możliwość integracji z systemami sterowania robotów.

### 2.3.3 System wspomagania nawigacji osób niewidomych

W artykule [3] poświęconym konferencji ICAAIC dotyczącej sztucznej inteligencji i rozumowania maszynowego (ang. *Applied Artificial Intelligence and Computing (ICAAIC)*) autorzy przedstawili system wspomagający nawigację osób niewidomych, wykorzystujący platformę Raspberry Pi oraz algorytm YOLO do detekcji obiektów. Głównym celem systemu jest ułatwienie osobom z dysfunkcją wzroku poruszania się w codziennym otoczeniu poprzez identyfikację i lokalizację przeszkód w czasie rzeczywistym.

**Komponenty systemu:**

- **Raspberry Pi:** Platforma obliczeniowa, na której uruchomiono algorytm YOLO i przetwarzanie danych z kamery.
- **Algorytm YOLO (ang. *You Only Look Once*):** Szybkie i dokładne wykrywanie obiektów w obrazie w czasie rzeczywistym.
- **Moduł kamery:** Rejestruje obraz otoczenia, który jest analizowany przez system.
- **Interfejs użytkownika:** Przekazuje informacje zwrotne w formie dźwiękowej lub wibracyjnej, informując użytkownika o wykrytych przeszkodach.

**Działanie systemu:** System analizuje obraz z kamery w czasie rzeczywistym, identyfikuje obiekty znajdujące się na drodze użytkownika i dostarcza odpowiednie informacje zwrotne. Użytkownik jest w ten sposób ostrzegany o przeszkodach, co umożliwia unikanie kolizji i bezpieczne poruszanie się.

**Zalety rozwiązań:**

- **Niski koszt:** Wykorzystanie Raspberry Pi oraz otwartoźródłowego oprogramowania czyni system przystępnym cenowo.
- **Mobilność:** Kompaktowe komponenty pozwalają na łatwe przenoszenie oraz integrację z akcesoriami, takimi jak laska dla niewidomych.
- **Skalowalność:** Możliwość rozbudowy systemu o dodatkowe funkcje, np. rozpoznanie znaków drogowych lub sygnalizacji świetlnej.

**Wnioski:** Przedstawiony system stanowi obiecujące narzędzie wspierające osoby niewidome w samodzielnym poruszaniu się, zwiększając ich bezpieczeństwo i komfort życia. Integracja technologii, takich jak Raspberry Pi oraz algorytmy głębokiego uczenia, umożliwia tworzenie efektywnych i dostępnych rozwiązań wspomagających nawigację.

$$y = \frac{\partial x}{\partial t} \quad (2.1)$$

jak i pojedyncze symbole  $x$  i  $y$  składa się w trybie matematycznym.

# Rozdział 3

## Przygotowanie i trening modelu detekcji obiektów

### 3.1 Przygotowanie danych wzorcowych

Do poprawnego wytrenowania sieci neuronowej, konieczne jest przygotowanie zestawu danych wzorcowych. W przypadku algorytmów głębokiego uczenia kluczowe znaczenie ma jakość i zróżnicowanie danych treningowych, ponieważ sieci neuronowe uczą się rozpoznawania wzorców na podstawie dostarczonych przykładów.

Dane wzorcowe muszą spełniać pewne istotne wymagania, aby zapewnić odpowiednią skuteczność modelu:

- **Różnorodność scenariuszy:** Obrazy powinny przedstawać obiekt w różnych warunkach oświetleniowych (np. światło dzienne, sztuczne oświetlenie, cienie), co pozwala modelowi uogólniać na nowe środowiska.
- **Zmienne pozycje i orientacje:** Dane powinny zawierać obiekt w różnych orientacjach, kątach oraz pozycjach względem kamery, aby uniknąć problemu nadmiernego dopasowania do specyficznego układu.
- **Różne odległości od kamery:** Obiekty powinny być widoczne w kadrze zarówno z bliska, jak i z daleka, co ułatwia modelowi detekcję niezależnie od dystansu.

W praktyce przygotowanie takiego zbioru danych wymaga użycia narzędzi do annotacji, które pozwalają na oznaczenie obiektów w obrazach za pomocą ramek ograniczających (ang. *bounding boxes*). Do tego celu często wykorzystuje się narzędzia do annotacji zdjęć, takie jak **Roboflow** czy **LabelImg**, które wspomagają proces podziału na zbiory treningowe, walidacyjne i testowe.

## 3.2 Platforma Roboflow

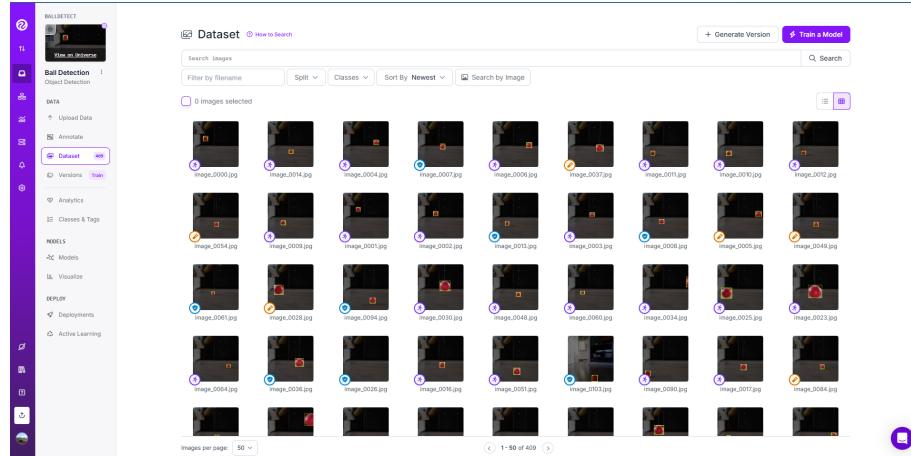
Aby zrealizować założenia pracy, do przygotowania zbioru uczącego wykorzystano platformę Roboflow. Jest to narzędzie dostępne online, które umożliwia zarządzanie danymi, ich etykietowanie oraz konwersję między różnymi formatami. W kontekście tej pracy wykorzystano obrazy przedstawiające czerwoną piłkę w realistycznym środowisku.

Proces przygotowania danych w Roboflow składał się z następujących etapów:

1. **Zbieranie danych** - zgromadzono zestaw około 400 zdjęć przedstawiających czerwoną piłkę w różnych warunkach oświetleniowych i na różnych tłaach.
2. **Annotacja** - każde zdjęcie zostało oznaczone poprzez narysowanie ramki ograniczającej wokół piłki. Roboflow udostępnia intuicyjny interfejs do tego procesu.
3. **Augmentacja danych** - aby zwiększyć różnorodność zbioru treningowego, zastosowano następujące techniki augmentacji:
  - Zmiana jasności
  - Rozmycie
  - Zróżnicowanie otoczenia
4. **Podział danych** - zbiór został podzielony w proporcjach:
  - 70% - zbiór treningowy
  - 20% - zbiór walidacyjny
  - 10% - zbiór testowy
5. **Eksport** - przygotowane dane zostały wyeksportowane w formacie YOLO v7 Pytorch.

### 3.2.1 Proces użycia witryny Roboflow

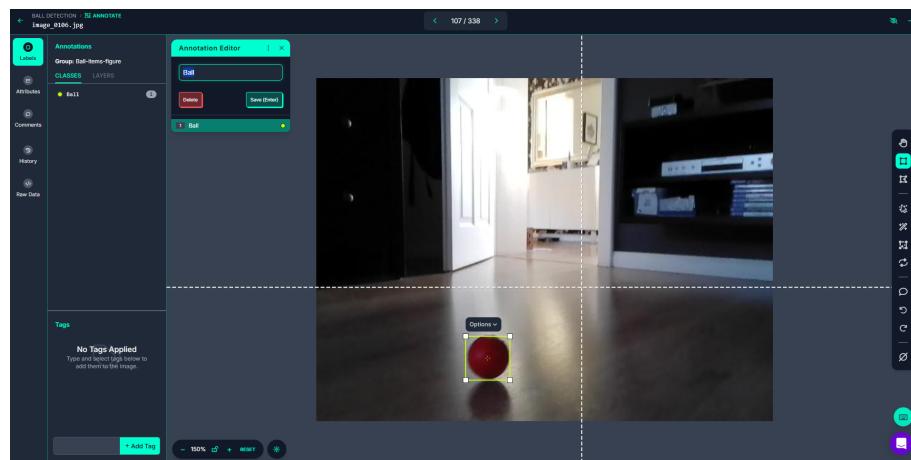
W celu przygotowania danych treningowych, zarejestrowano konto na platformie Roboflow i przystąpiono do etykietowania obiektów. Na rysunku 3.1 przedstawiono widok strony głównej witryny Roboflow.



Rysunek 3.1: Widok strony głównej witryny Roboflow.

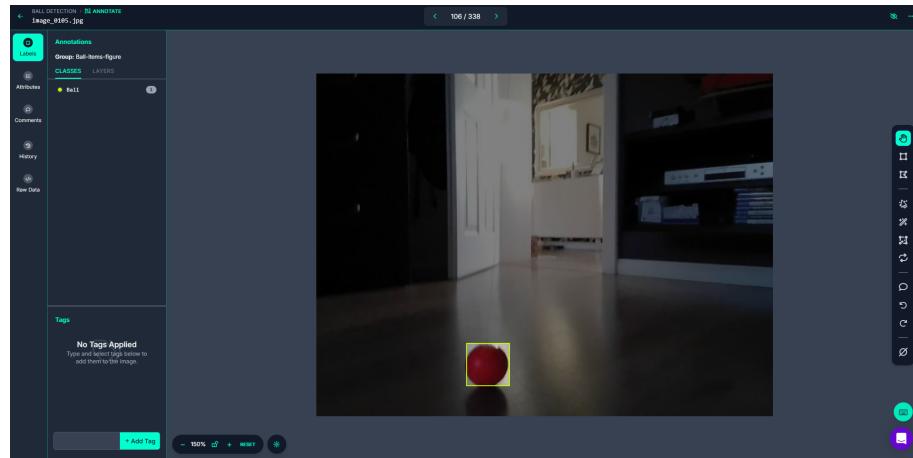
Proces etykietowania można przeprowadzić automatycznie, korzystając z algorytmów detekcji obiektów, lub ręcznie, poprzez zaznaczenie obiektów na obrazie. Korzystając z etykietowania ręcznego istnieje pewność, że obiekty zostaną poprawnie oznaczone w takich przypadkach jak rozmyty obraz bądź słabe oświetlenie.

**Zdjęcie 3.2** przedstawia interfejs użytkownika podczas etykietowania obiektów. Na obrazie widoczna jest czerwona piłka, która została zaznaczona ramką ograniczającą.



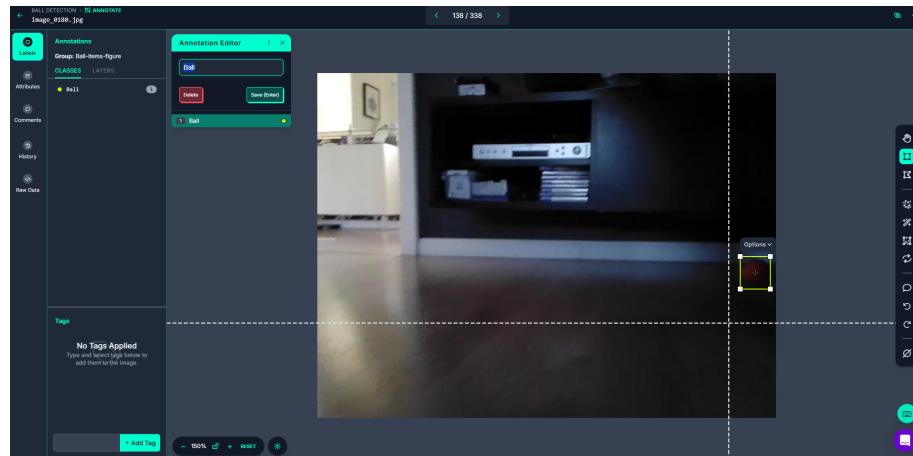
Rysunek 3.2: Etykietowanie obiektów za pomocą witryny Roboflow.

**Zdjęcie 3.3** przedstawia obiekt po zetykietowaniu. Na obrazie widoczna jest czerwona piłka, która została poprawnie oznaczona ramką ograniczającą.



Rysunek 3.3: Obiekt po zetykietowaniu.

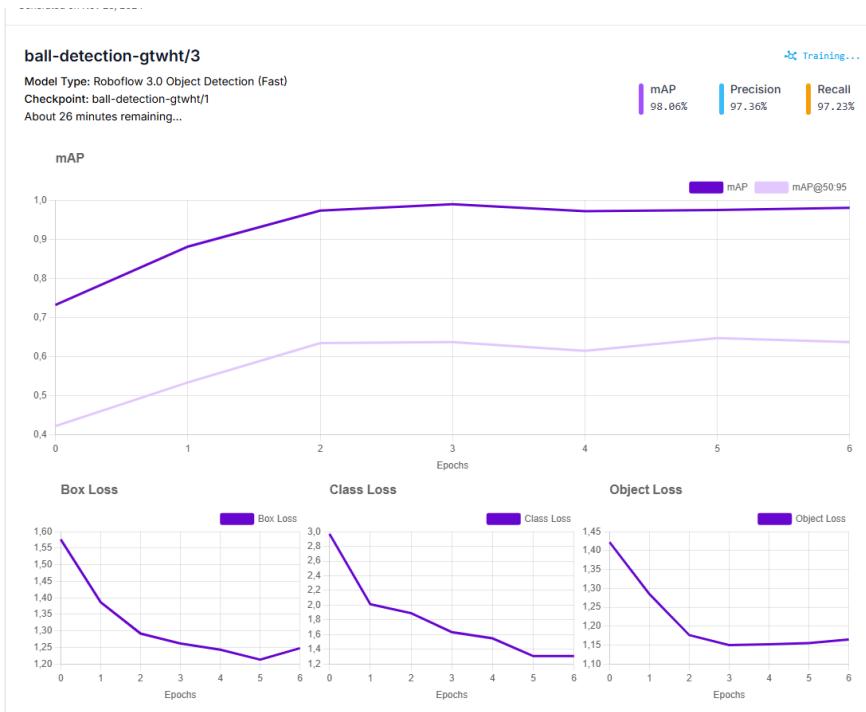
**Zdjęcie 3.4** pokazuje proces etykietowania w warunkach słabego oświetlenia. Pomimo trudnych warunków, obiekt został poprawnie oznaczony, co świadczy o zalecie ręcznego etykietowania w takich przypadkach.



Rysunek 3.4: Etykietowanie w słabym oświetleniu.

### 3.2.2 Trening zestawu danych

Po zakończeniu etykietowania danych, przystąpiono do segmentacji zbioru na zbiór treningowy, walidacyjny i testowy. Następnie przeprowadzono trening danych aby można było wyeksportować je w formacie YOLO v7 Pytorch. Na wykresie 3.5 przedstawiono przebieg treningu w pierwszych epokach. Widać, że wartość funkcji straty maleje wraz z postępem uczenia, co świadczy o skuteczności procesu.



Rysunek 3.5: Wykres pierwszych epok treningu.

Natomiast wykres 3.6 przedstawia przebieg treningu w kolejnych epokach. Widać, że w momencie 20 epoki wartość wskaźnika mAP maleje. **mAP (mean Average Precision)** jest miarą jakości modeli detekcji obiektów, uwzględniającą zarówno *precision*, jak i *recall*. Oblicza się go jako średnią precyzji (AP) dla wszystkich klas obiektów.

Podczas treningu mAP może zmaleć, a następnie wzrosnąć z kilku powodów:

- **Augmentacja danych:** Trudniejsze przypadki augmentacji danych mogą chwilowo obniżyć mAP, ale poprawiają generalizację modelu.
- **Dane walidacyjne:** Specyficzne, trudne obrazy w danych walidacyjnych mogą wpływać na chwilowe obniżenie wyników.

Spadek mAP jest naturalnym zjawiskiem, oznaczającym reakcję modelu na zmiany w procesie treningu. Wzrost mAP po spadku wskazuje na poprawę generalizacji oraz stabilizację procesu optymalizacji.



Rysunek 3.6: Wykres przebiegu treningu z zaistniałym spadkiem mAP.

### 3.3 Trening modelu YOLOv7

Po przygotowaniu danych treningowych przystąpiono do wytrenowania modelu detekcji obiektów YOLOv7, korzystając z platformy Google Colab. Umożliwia ona wykorzystanie zasobów obliczeniowych w chmurze, w tym akceleracji przez GPU (Graphics Processing Unit), które znacząco przyspiesza przetwarzanie dużych ilości danych dzięki równoległemu wykonywaniu obliczeń. Jest to szczególnie istotne w przypadku modeli uczenia maszynowego, takich jak głębokie sieci neuronowe, które wymagają intensywnych operacji macierzowych. Proces treningu przebiegał zgodnie z następującymi krokami:

#### 3.3.1 Przygotowanie środowiska

Utworzenie nowego projektu w Google Colab i skonfigurowano środowiska do korzystania z GPU:

1. Z paska narzędzi **Środowisko wykonawcze** wybrano opcję **Zmień typ środowiska wykonawczego**.
2. Jako **Akcelerator sprzętowy** ustawiono GPU.

#### 3.3.2 Instalacja wymaganych bibliotek

Pobranie repozytorium YOLOv7 z platformy GitHub oraz zainstalowano niezbędne zależności:

```
!git clone https://github.com/WongKinYiu/yolov7  
%cd yolov7  
!pip install -r requirements.txt
```

#### 3.3.3 Pobranie i przygotowanie danych treningowych

Dane treningowe zostały przygotowane na platformie Roboflow, gdzie wygenerowano klucz API, który umożliwił automatyczne pobranie danych do środowiska Google Colab. W notebooku dodano poniższy fragment kodu do zainportowania danych:

```
!pip install roboflow  
from roboflow import Roboflow  
rf = Roboflow(api_key="TWÓJ_KLUCZ_API")  
project = rf.workspace().project("nazwa-projektu")  
dataset = project.version(1).download("yolov7")
```

### 3.3.4 Przygotowanie modelu

Pobranie wstępnie wytrenowanych wag modelu YOLOv7:

```
!wget https://github.com/WongKinYiu/yolov7/releases/download/v0.1/yolov7.pt
```

### 3.3.5 Trening modelu

Model wytrenowano, uruchamiając poniższą komendę:

```
%env WANDB_MODE=disabled  
!python yolov7/train.py --data Ball-Detection-3/data.yaml --cfg yolov7/cfg/training/yolov7.pt
```

Parametry treningu były następujące:

- **-data**: Ścieżka do pliku konfiguracyjnego zestawu danych otrzymanych z Roboflow.
- **-cfg**: Plik konfiguracyjny architektury modelu.
- **-weights**: Wstępnie wytrenowane wagi modelu YOLOv7.
- **-epochs**: Liczba epok (ustawiono na 50).
- **-batch-size**: Rozmiar grupy treningowej (ustawiono na 16).
- **-device**: GPU (0 oznacza pierwsze dostępne GPU).

Po zakończeniu treningu najlepsze wagi modelu są zapisywane w pliku **best.pt**. Plik ten będzie wykorzystywany w programie do detekcji obiektu w czasie rzeczywistym.

# Rozdział 4

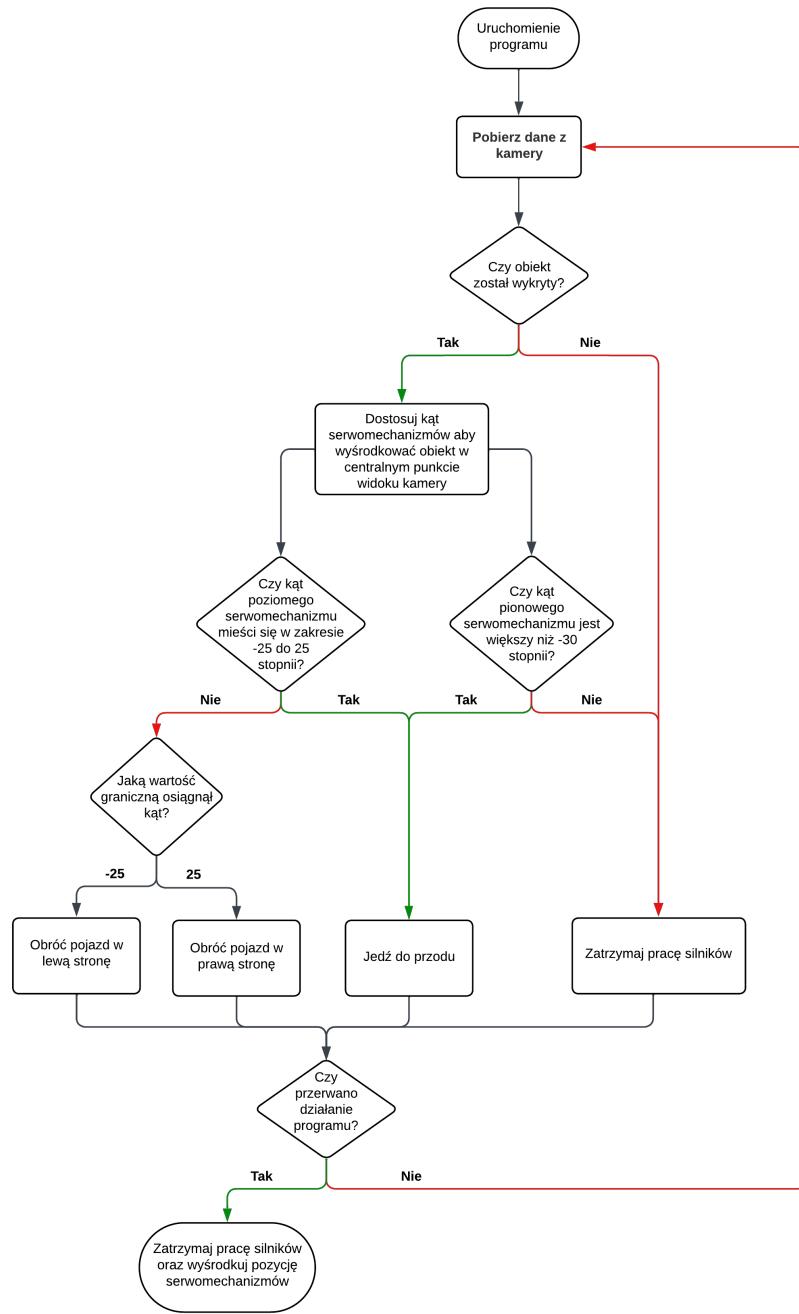
## Proces tworzenia robota mobilnego

W obecnym rozdziale przedstawiony zostanie proces tworzenia robota mobilnego, uwzględniający zarówno założenia logiczne, jak i aspekty techniczne związane z budową oraz implementacją systemów sterowania. W pierwszej części opisane zostaną kluczowe założenia projektowe oraz schemat blokowy algorytmu robota mobilnego opartego na systemie wizyjnym. Następnie zostanie zaprezentowany proces konstrukcji robota, opierając się na modelu 3D stworzonym w programie Autodesk Fusion oraz przedstawienie wykorzystanych modułów i ich sposób podłączenia przedstawionego za pośrednictwem schematu blokowego. Ostatnia część rozdziału dotyczy dostosowania jednostki sterującej w postaci Raspberry Pi, w tym konfiguracji niezbędnych bibliotek umożliwiających integrację podzespołów i realizację funkcjonalności systemu wizyjnego dla robota mobilnego.

### 4.1 Założenia logiczne

Poniżej przedstawiono diagram blokowy ilustrujący działanie algorytmu detekcji obiektów w czasie rzeczywistym. Proces rozpoczyna się od akwizycji obrazu z kamery, który następnie trafia do modelu detekcji obiektów YOLOv7. Na podstawie wyników detekcji robot dostosowuje kąty serwomechanizmów pionowego i poziomego, aby utrzymać obiekt w centrum kadru. Jeśli obiekt przesunie się poza wyznaczone centrum, program oblicza różnicę w pikselach i przelicza ją na odpowiednie wartości kątowe, umożliwiając sterowanie serwomechanizmami w celu zapewnienia ciągłości śledzenia. Celem algorytmu jest nieustanne utrzymanie obiektu w centrum widoku kamery. Program na bieżąco monituruje wartości kątowe serwomechanizmów, które są wykorzystywane do sterowania pracą silników DC odpowiedzialnych za ruch robota. Ruch w przód wykonywany jest jedynie w przypadku spełnienia następujących warunków: obiekt został wykryty, kąt serwomechanizmu poziomego mieści się w zakresie od  $-25^\circ$  do  $25^\circ$ , a kąt serwomechanizmu pionowego przekracza  $-30^\circ$ . W przeciwnym razie robot zatrzymuje się lub wykonuje skręt. Jeśli kąt poziomy osiągnie wartość graniczną ( $25^\circ$  lub  $-25^\circ$ ), robot skręca odpowiednio w prawo lub w lewo.

Program można zatrzymać wyłącznie przez interakcję użytkownika, co powoduje zatrzymanie wszystkich systemów i powrót robota do stanu początkowego. Stan początkowy definiuje sytuację, w której serwomechanizmy ustawione są w pozycji centralnej ( $0^\circ$  dla obu osi), a silniki DC są wyłączone.



Rysunek 4.1: Schemat blokowy działania algorytmu pracy robota z systemem wizyjnym

## 4.2 Konstrukcja robota mobilnego

Konstrukcja robota mobilnego oparta jest na platformie jezdnej z dwoma silnikami DC o napędzie różnicowym, serwomechanizmem pionowym i poziomym oraz kamerą. Robot wyposażony jest w moduł Raspberry Pi 4B, który pełni rolę jednostki decyzyjnej oraz komunikacyjnej. Poniżej przedstawiono schemat blokowy robota mobilnego z zaznaczonymi podstawowymi elementami składowymi.

### 4.2.1 Wykorzystane moduły

Aby zrealizować założenia projektowe, wykorzystano następujące moduły i komponenty:

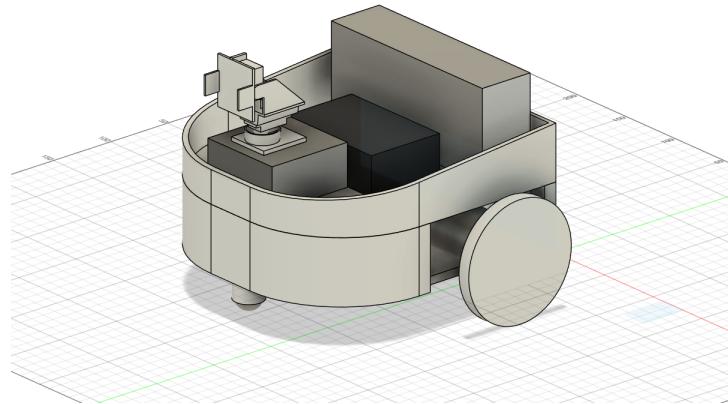
- **Raspberry Pi:** Raspberry Pi 4B WiFi 8GB RAM - jednostka centralna robota, odpowiedzialna za przetwarzanie obrazu, sterowanie silnikami oraz serwomechanizmami.
- **Kamera:** Raspberry Pi Camera HD v3 12MPx, służąca do akwizycji obrazu.
- **Moduł sterownika silników:** Cytron HAT-MDD10 - dwukanałowy sterownik silników DC 24V/10A z funkcją PWM służący jako nakładka na Raspberry Pi.
- **Silniki DC:** Dwa silniki z przekładnią 37Dx68L 30:1 12V 330RPM + enkoder CPR 64 - Pololu 4752 - napędzające koła robota.
- **Zasilanie:** Cztery ogniska 18650 Li-Ion Samsung INR18650-35E 3500mAh 3,6V - zasilanie silników DC.
- **Serwomechanizm pionowy:** Serwo TowerPro SG-90 micro 180 stopni - odpowiada za regulację kąta nachylenia kamery w pionie.
- **Serwomechanizm poziomy:** Serwo TowerPro SG-92R micro - odpowiada za regulację kąta obrotu kamery w poziomie.
- **Uchwyt kamery:** Wykonany z połączenia gotowego uchwytu do serw micro Pan/Tilt oraz dedykowanego uchwytu do kamery Raspberry Pi v2 oraz v3.
- **Obudowa:** Wykonana przy pomocy drukarki 3D z wytrzymałego filamentu Spectrum PETG 1,75mm.
- **Koła:** Dwa Koła Pololu 90x10mm - Zapewniające ruch robota po powierzchni.
- **Zasilanie Raspberry Pi:** Powerbank o parametrach 30000mAh 5V 3A.
- **Mocowanie silników:** Dwie sztuki mocowań aluminiowych do silników 37D Pololu 1084.

- **Huby mocujące:** Dwa aluminiowe huby mocujące 6mm M3 do kół Pololu 1999.
- **Koło wspornikowe pasywne:** Ball Caster 3/4" metalowe Pololu 955 - zapewniające stabilność robota.
- **Taśma do kamery:** Taśma Raspberry Pi kamera 60cm 15 żyłowa raster 1mm - łącząca kamerę z Raspberry Pi.
- **Przewody:** Zestaw przewodów połączeniowych justPi męsko-męskich, żeńsko-żeńskich, żeńsko-męskich.
- **Przełącznik:** Przełącznik dźwigniowy ON-OFF KN3(C)-101 250V/6A - służący do ręcznego wyłączania zasilania silników DC.

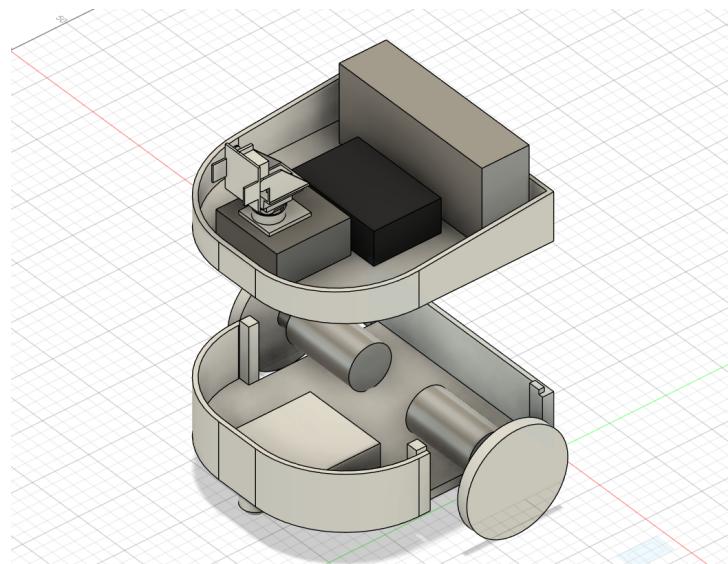
### 4.2.2 Schemat elektryczny - dokonczyc

### 4.2.3 Projekt robota mobilnego

Jednym z założeń projektu było stworzenie kompaktowego robota mobilnego o możliwie prostym i modułowym układzie konstrukcyjnym. Robot składa się z dwóch pięter: dolnego pełniącego funkcję napędową oraz górnego będącego platformą dla modułów sterujących. Dzięki temu podziałowi możliwe jest zachowanie walorów estetycznych oraz funkcjonalnych.

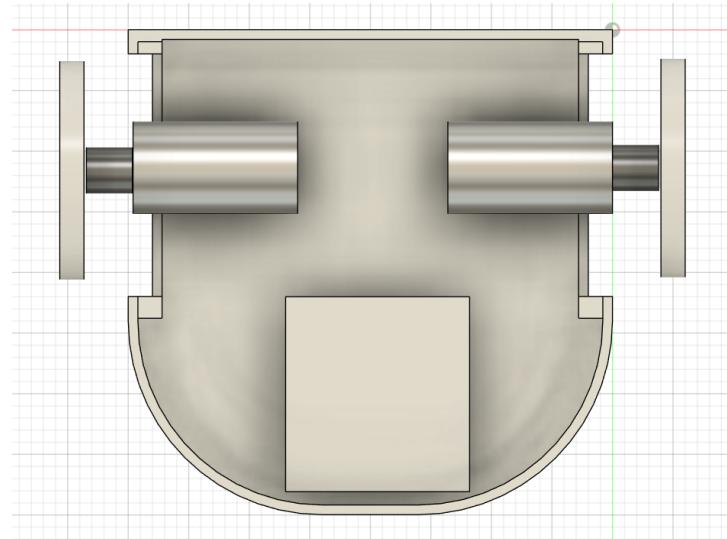


Rysunek 4.2: Projekt konstrukcji robota mobilnego wykonany w programie Autodesk Fusion.



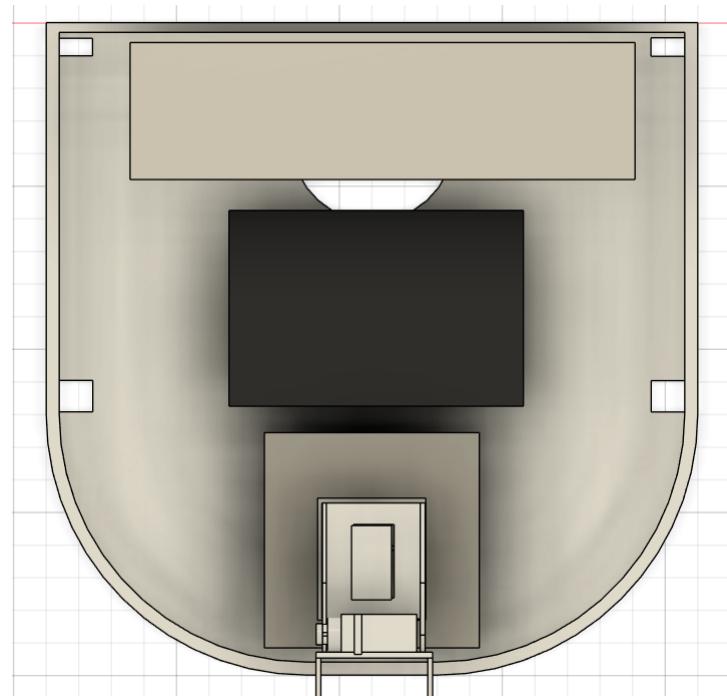
Rysunek 4.3: Podział konstrukcji na piętra.

Obraz 4.4 przedstawia dolne piętro robota, na którym zamontowane są silniki DC wraz z kołami oraz zasilaniem w postaci koszyka z czterema ogniwami 18650 Litowo-jonowymi. Dodatkowo umieszczono przełącznik dźwigniowy ON-OFF, który umożliwia ręczne wyłączenie zasilania silników DC.



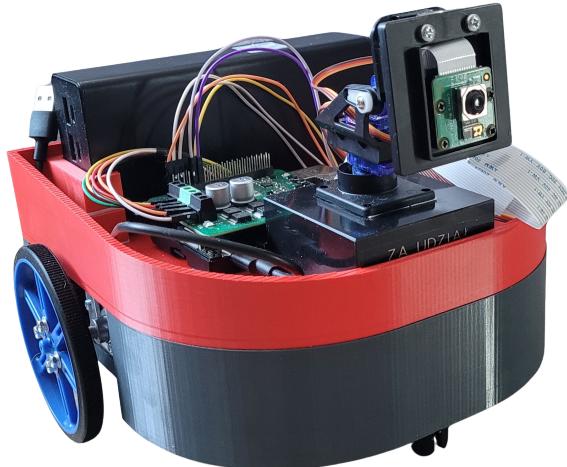
Rysunek 4.4: Widok rzutu górnego na dolne piętro robota.

Obraz 4.5 przedstawia górne piętro robota, na którym zamontowane są takie moduły jak Raspberry Pi, sterownik silników DC, uchwyt na kamerę wraz z serwomechanizmami oraz zasilaniem. Z powodu dużej wagi powerbanka, konieczne było zastosowanie przeciwagi w postaci granitowego bloku na przodzie modułu górnego piętra. Zastosowanie to zapewnia stabilność robota oraz równomierny rozkład masy.



Rysunek 4.5: Widok rzutu górnego na górne piętro robota.

Zdjęcie 4.6 przedstawia gotową konstrukcję robota mobilnego wykonanego z tworzywa sztucznego PETG.



Rysunek 4.6: Gotowa konstrukcja robota mobilnego.

## 4.3 Dostosowanie RaspberryPi

Na Raspberry Pi zainstalowano system operacyjny Debian GNU/Linux 12 (Bookworm), który jest 64-bitową wersją opartą na jądrze Linux 6.6.51, zoptymalizowaną pod architekturę ARM (aarch64). Wybrana wersja systemu jest kompatybilna z bibliotekami istotnymi do poprawnego działania systemu wizyjnego. Aby zoptymalizować pracę programu i zapewnić pełną kontrolę nad jego zależnościami, utworzono dedykowane wirtualne środowisko, w którym przechowywane są wszystkie wymagane biblioteki. Środowisko to zostało stworzone przy użyciu narzędzia `venv` za pomocą poniższego polecenia w terminalu:

```
python3 -m venv nazwa-env
```

Po utworzeniu środowiska wirtualnego należy je aktywować, używając polecenia:

```
source nazwa-env/bin/activate
```

Program jest uruchamiany wewnątrz tego środowiska, co pozwala na uniknięcie konfliktów między wersjami bibliotek oraz zapewnia powtarzalność działania systemu na różnych urządzeniach. Przykładowe polecenie uruchamiające program w aktywnym środowisku to:

```
python3 program.py
```

Takie podejście gwarantuje, że wszystkie zależności programu są odpowiednio załadowane i kompatybilne z systemem operacyjnym oraz sprzętem Raspberry Pi.

### 4.3.1 Spis wymaganych bibliotek

Poniżej przedstawiono listę bibliotek wymaganych do uruchomienia systemu wizyjnego oraz sterowania robotem mobilnym, z podziałem na ich funkcje:

- **Platformy uczenia maszynowego:**
  - **torch** – Podstawowa platforma do trenowania i uruchamiania modeli uczenia maszynowego.
  - **models.experimental** – Obsługa modeli YOLO do detekcji obiektów.
  - **utils.general** – Funkcje wspomagające dla YOLO, takie jak filtrowanie wyników i skalowanie współrzędnych.
- **Przetwarzanie obrazu:**
  - **opencv-python (cv2)** – Obsługa obrazów i przetwarzanie wizyjne, np. binaryzacja i filtrowanie.
  - **numpy** – Operacje matematyczne i macierzowe, kluczowe dla analizy danych wizualnych.
  - **Picamera2** – Obsługa kamery Raspberry Pi do akwizycji obrazu.
- **Sterowanie sprzętem:**
  - **RPi.GPIO** – Sterowanie pinami GPIO na Raspberry Pi.
  - **pigpio** – Niskopoziomowe sterowanie GPIO, np. dla serwomechanizmów.
- **Regulator PID:**
  - **simple\_pid** – Implementacja regulatora PID, używana do precyzyjnego sterowania, np. serwomechanizmami.
- **Inne ważne biblioteki:**
  - **tqdm** – Pasek postępu przy iteracjach.
  - **PyYAML** – Obsługa plików konfiguracji w formacie YAML.

Szczegółowy spis wszystkich bibliotek został przedstawiony w pliku Requirements.txt dostarczonym w załączniku do pracy.



# Rozdział 5

## Analiza działania systemu wizyjnego

W niniejszym rozdziale przedstawiono analizę działania systemu wizyjnego, który został zaimplementowany w ramach pracy. System ten umożliwia detekcję obiektu na obrazie z kamery Raspberry Pi, a następnie sterowanie serwomechanizmami w celu utrzymania obiektu w centrum kadru. W ramach analizy porównano skuteczność dwóch algorytmów detekcji obiektów: YOLOv7 oraz binaryzację i klasyczne techniki przetwarzania obrazów. W celu oceny skuteczności algorytmów przeprowadzono testy na zestawie danych treningowych, a następnie porównano wyniki. W kolejnych sekcjach przedstawiono opis działania programu, uzyskane rezultaty oraz porównanie skuteczności algorytmów.

### 5.1 Detekcja obiektu za pomocą narzędzi głębokiego uczenia

#### 5.1.1 Opis działania programu

---

```
1   class test : public basic
2   {
3       public:
4           test (int a);
5           friend std::ostream operator<<(std::ostream & s,
6                                           const test & t);
7       protected:
8           int _a;
9
10 };
```

---

Rysunek 5.1: Pseudokod w listings.

### **5.1.2 Uzyskane rezultaty**

## **5.2 Detekcja obiektu za pomocą binaryzacji i klasycznych technik przetwarzania obrazów**

## **5.3 Porównanie skuteczności algorytmów**

## Rozdział 6

# Problematyka i rozwiązania



# Rozdział 7

## Podsumowanie i wnioski

- sposób testowania w ramach pracy (np. odniesienie do modelu V)
- organizacja eksperymentów
- przypadki testowe zakres testowania (pełny/niepełny)
- wykryte i usunięte błędy
- opcjonalnie wyniki badań eksperimentalnych

Tabela 7.1: Nagłówek tabeli jest nad tabelą.

$\zeta$	metoda							
	alg. 1	alg. 2	alg. 3			alg. 4, $\gamma = 2$		
			$\alpha = 1.5$	$\alpha = 2$	$\alpha = 3$	$\beta = 0.1$	$\beta = -0.1$	
0	8.3250	1.45305	7.5791	14.8517	20.0028	1.16396	1.1365	
5	0.6111	2.27126	6.9952	13.8560	18.6064	1.18659	1.1630	
10	11.6126	2.69218	6.2520	12.5202	16.8278	1.23180	1.2045	
15	0.5665	2.95046	5.7753	11.4588	15.4837	1.25131	1.2614	
20	15.8728	3.07225	5.3071	10.3935	13.8738	1.25307	1.2217	
25	0.9791	3.19034	5.4575	9.9533	13.0721	1.27104	1.2640	
30	2.0228	3.27474	5.7461	9.7164	12.2637	1.33404	1.3209	
35	13.4210	3.36086	6.6735	10.0442	12.0270	1.35385	1.3059	
40	13.2226	3.36420	7.7248	10.4495	12.0379	1.34919	1.2768	
45	12.8445	3.47436	8.5539	10.8552	12.2773	1.42303	1.4362	
50	12.9245	3.58228	9.2702	11.2183	12.3990	1.40922	1.3724	

# Bibliografia

- [1] Przemysław Łabęcki i Andrzej Kasiński. „Aktywny pokładowy system wizyjny dla robota autonomicznego”. W: *Pomiary Automatyka Kontrola* 55.9 (2009), s. 731–736.
- [2] Kamil Miękus, Tomasz Marek, Paweł Proszkowiec i Jerzy Zając. „System wizyjny 3D kołowego robota mobilnego”. W: *Logistyka* 6 (2014), s. 7442–7442.
- [3] R Parvadhavardhni, Pankhuri Santoshi i A Mary Posonia. „Blind navigation support system using Raspberry Pi & YOLO”. W: *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*. IEEE. 2023, s. 1323–1329.
- [4] Joseph Redmon, Santosh Divvala, Ross Girshick i Ali Farhadi. „You Only Look Once: Unified, Real-Time Object Detection”. W: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016* (2016). DOI: 10.1109/CVPR.2016.91, s. 779–788. URL: <https://arxiv.org/abs/1506.02640>.
- [5] Chien-Yao Wang, Alexey Bochkovskiy i Hong-Yuan Mark Liao. „YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors”. W: *arXiv preprint arXiv:2207.02696* (2022). Available online: <https://arxiv.org/abs/2207.02696>. URL: <https://arxiv.org/abs/2207.02696>.



## **Dodatki**



# Spis skrótów i symboli

DNA kwas deoksyrybonukleinowy (ang. *deoxyribonucleic acid*)

MVC model – widok – kontroler (ang. *model–view–controller*)

$N$  liczebność zbioru danych

$\mu$  stopnień przyleżności do zbioru

$\mathbb{E}$  zbiór krawędzi grafu

$\mathcal{L}$  transformata Laplace'a



# Źródła

Jeżeli w pracy konieczne jest umieszczenie długich fragmentów kodu źródłowego, należy je przenieść w to miejsce.

---

```
1 if (_nClusters < 1)
2     throw std::string ("unknown number of clusters");
3 if (_nIterations < 1 and _epsilon < 0)
4     throw std::string ("You should set a maximal number of
5         iteration or minimal difference --- epsilon .");
6 if (_nIterations > 0 and _epsilon > 0)
7     throw std::string ("Both number of iterations and minimal
8         epsilon set --- you should set either number of iterations
9         or minimal epsilon .");
```

---



# **Lista dodatkowych plików, uzupełniających tekst pracy**

W systemie do pracy dołączono dodatkowe pliki zawierające:

- źródła programu,
- dane testowe,
- film pokazujący działanie opracowanego oprogramowania lub zaprojektowanego i wykonanego urządzenia,
- itp.



# Spis rysunków

3.1	Widok strony głównej witryny Roboflow. . . . .	13
3.2	Etykietowanie obiektów za pomocą witryny Roboflow. . . . .	13
3.3	Obiekt po zetykietowaniu. . . . .	14
3.4	Etykietowanie w słabym oświetleniu. . . . .	14
3.5	Wykres pierwszych epok treningu. . . . .	15
3.6	Wykres przebiegu treningu z zaistniałym spadkiem mAP. . . . .	16
4.1	Schemat blokowy działania algorytmu pracy robota z systemem wizyjnym .	20
4.2	Projekt konstrukcji robota mobilnego wykonany w programie Autodesk Fusion. . . . .	23
4.3	Podział konstrukcji na piętra. . . . .	23
4.4	Widok rzutu górnego na dolne piętro robota. . . . .	24
4.5	Widok rzutu górnego na górne piętro robota. . . . .	24
4.6	Gotowa konstrukcja robota mobilnego. . . . .	25
5.1	Pseudokod w listings. . . . .	29



# **Spis tabel**

7.1 Nagłówek tabeli jest nad tabelą.	34
--------------------------------------	----