

Zaawansowane Przygotowanie Danych w Uczniu Maszynowym  
Systemy z Uczniem Maszynowym

# Klasyfikacja piłek sportowych na podstawie zdjęć

## RAPORT 1

Norbert Nieżorawski 180050

Przemysław Rośleń 180150

Aleksandra Rumińska 179953

Wiktor Krasiński 179987

Jakub Nowak 180503

---

### Definicja problemu

Zadaniem systemu będzie klasyfikacja piłek sportowych ze zdjęć zrobionych telefonem komórkowym. Uwzględnione zostaną piłki z różnych dyscyplin sportowych takich jak: tenis, koszykówka, piłka nożna, tenis stołowy, siatkówka oraz piłka ręczna. Wszystkie zdjęcia będą zawierały tylko jedną piłkę danego rodzaju, gdyż w przypadku wystąpienia wielu piłek jednego rodzaju lub wielu piłek różnego rodzaju, algorytm mógłby mieć duże trudności z klasyfikacją. Ponadto zdjęcia będą wykonane na zróżnicowanym tle i przy różnorodnym oświetleniu, aby zapobiec sytuacji, w której model umiałby klasyfikować piłki tylko występujące na danym tle, a nie radził sobie przy innym. Będzie to dla modelu wyzwanie, którego rozwiązanie pozwoli na rozpoznawanie właściwych obiektów niezależnie od ich otoczenia, przez co zwiększy się jego generalizacja. W przypadku, gdy na zdjęciu nie zostanie wykryta żadna piłka, model oznaczy obraz etykietą "inne".

### Wejście modelu

Do wejścia modelu zostaną wykorzystane obrazy zrobione telefonem komórkowym. Ich formatem będzie RGB w rozszerzeniu z JPEG, gdyż są to domyślne ustawienia zdjęć zrobionych takim sprzętem. By poprawić jakość obrazów i zredukować wpływ czynników takich jak oświetlenie, tło czy rozmiar zdjęcia, może być konieczne dokonanie obróbki przed przetworzeniem obrazu przez model.

Przykłady danych wejściowych:



## Wyjście modelu

Zadaniem modelu jest przyporządkowanie każdego zdjęcia piłki do jednej z sześciu klas reprezentujących różne dyscypliny sportowe w ramach zadania klasyfikacji. Następnie przypisane przez model etykiety zostaną porównane z etykietami, które zostały wprowadzone ręcznie do zbioru testowego we wcześniejszym etapie. W przypadku, gdy model nie wykryje piłki na zdjęciu, zostanie ona przypisana do klasy "inne".

Przykładowe wyjścia dla przykładowych wejść:

	piłka do tenisa
	piłka do koszykówki

	piłka do nożnej
	piłeczka do tenisa stołowego
	piłka do siatkówki
	piłka do ręcznej
	inne

Zaawansowane Przygotowanie Danych w Uczniu Maszynowym  
Systemy z Uczniem Maszynowym

# Klasyfikacja piłek sportowych na podstawie zdjęć

## RAPORT 2

Norbert Nieżorawski 180050

Przemysław Rośleń 180150

Aleksandra Rumińska 179953

Wiktor Krasiński 179987

Jakub Nowak 180503

---

### 1. Opisz dane, które są potrzebne w celu treningu i testowania

W naszym projekcie będziemy potrzebowali zdjęć piłek sportowych z dyscyplin takich jak: koszykówka, siatkówka, piłka nożna, tenis, piłka ręczna, tenis stołowy. Pozyskamy te dane ze sklepów ze sprzętem sportowym, z prywatnych zasobów przy pomocy własnych telefonów komórkowych oraz z publicznie dostępnych zbiorów danych znajdujących się na stronach internetowych:

- <https://www.dreamstime.com>
- <https://www.kaggle.com>
  - Sports balls - multiclass image classification
  - 30 Types of Balls Updated - Image Classification
- <https://www.istockphoto.com>

Każde ze zdjęć powinno być kolorowe i w formacie .jpg. Zdjęcia powinny posiadać różne rodzaje tła, różne oświetlenie, różne perspektywy. Piłki powinny być dobrze widoczne i znajdująć się w centralnym punkcie ekranu. Zdjęcia powinny mieć rozdzielcość 224x224, ponieważ jest to wystarczająco wysoka rozdzielcość, aby umożliwić sieciom neuronowym nauczenie się istotnych cech obiektów na zdjęciach, a jednocześnie wystarczająco niska, aby nie przeciążyć pamięci i mocy obliczeniowej modelu.

W fazie przygotowania danych, zdjęcia zostaną podzielone w różnych konfiguracjach na zbiory:

- treningowy i testowy
- treningowy, walidacyjny i testowy

Każdy ze zbiorów będzie zawierał zdjęcia piłek każdego rodzaju, zarówno zrobione przez nas, jak i pozyskane ze zbiorów dostępnych w internecie.

## 2. Zaplanuj procedurę zbierania danych

- a. Spróbuj przewidzieć, jakiego rodzaju dane będą potrzebne i jakie są potencjalne źródła danych

Zdjęcia kolorowe

Z telefonu

Orientacja pozioma

Różne przybliżenie

3:4 format

DECATHLON + WŁASNE PIŁKI

Podłoga w sklepie

Trawa

Stół do tenisa stołowego

- b. Spróbuj oszacować ilość danych, których będziesz potrzebować dla każdej klasy.

540 fotek

180 na osobę

30 na klasę <- 30 różnych piłek z jednej klasy

- c. Spróbuj przewidzieć potencjalne problemy związane z gromadzeniem danych, opisz sposoby minimalizowania tego ryzyka

tła mogą być zbyt podobne <- poświęcić więcej czasu na znalezienie odpowiedniego tła  
może nie być tyle piłek <- zdjęcia niektórych piłek będą zrobione więcej niż raz z różnym jej ułożeniem i kątem robienia zdjęć

różne egzemplarze piłek tego samego typu mogą być bardzo podobne do siebie (np. tenis, pingpong)

piłki będą zapakowane w zestawie i nie można ich rozpakować <- może trzeba będzie kupić piłki

problemy z oświetleniem na niektórych zdjęciach <- zrobić zdjęcia testowe/kilka zdjęć jednej piłki, aby wybrać to najlepsze//kontrolować jasność i ustawienia aparatu

problemy z jakością niektórych zdjęć <- będą wymagały dodatkowego przetworzenia, aby jakość była odpowiednia do treningu

może nie być tła, które sobie założyliśmy (np. stół do tenisa, trawa) <- znaleźć coś podobnego/zdjęcie stołu stołowego i na tym zrobić zdj/pójść w inne miejsce.

problemy z obsługą, ludzie będą się dziwnie patrzeć i nas wyprosić (podejrzana kradzież) <- pójść do innego sklepu

zrobione zdjęcie ucina kawałek piłki <- zrobić nowe zdjęcie i zastąpić je starym

inne przedmioty w sklepie mogą wejść w kadr <- znaleźć kąt, gdzie będą najmniej przeszkadzały/ gdzie będą najmniej widoczne  
różne sprzęty do robienia zdjęć, różne rozdzielczości aparatów członków zespołu <- potrzeba dodatkowego przetworzenia danych

### **3. Zbierz dane stosując powyższą procedurę. Zaprezentuj kilka przykładów i opisz problemy, jakie napotkałeś**

Udało nam się zebrać 540 własnych zdjęć, które zostały wykonane przez wszystkich członków zespołu (ZPD):

- Norbert Nieżorawski (2 prywatne telefony komórkowe)
- Aleksandra Rumińska (własny telefon komórkowy)
- Przemysław Rośleń (własny telefon komórkowy)

Podczas zbierania danych napotkaliśmy następujące problemy:

- Ciężko było znaleźć stół ping-pongowy jako tło do zbierania danych
- Aby dotrzeć do stołu do ping-ponga by zrobić zdjęcia trzeba było przetransportować niektóre piłki przez pół sklepu
- Oświetlenie w sklepie padające z różnych źródeł na suficie powodowały tworzenie się cieni.
- Podczas ustalania procedury zbierania zdjęć, nie zostało jasno określone, jak zdjęcia powinny być robione na stole ping-pongowym, co spowodowało, że powstały takie, które: mają tylko czarną powierzchnię stołu jako tło, uwzględniają namalowane na stole linie, uwzględniają siatkę znajdującą się na stole.
- Niektórych rodzajów piłek było bardzo mało (tylko 2-5 piłki do piłki ręcznej w zależności od sklepu)
- Niektóre piłki znajdowały się w opakowaniach, które trzeba było wyjąć (piłka do ping-ponga, piłka do tenisa)
- Niektóre rodzaje piłek pochodziły tylko od jednego producenta, przez co były bardzo do siebie podobne
- Piłki w sklepie były na bieżąco kupowane przez klientów, przez co zdarzało się, że piłka, której chcieliśmy zrobić zdjęcie została zabrana zanim je zrobiliśmy.
- Podczas robienia zdjęć na powierzchni trawo-podobnej w sklepie Decathlon, ciężko było znaleźć idealne miejsce, gdyż znajdowało się na niej wiele różnych sprzętów
- Jedna osoba nie zastosowała się do instrukcji wskazanej w punkcie 2 a) mówiącego o proporcjach zdjęć 3:4, przez co niektóre zdjęcia mają inne proporcje
- Po przeglądzie zdjęć, okazało się, że niektórzy członkowie drużyny nie wstawili odpowiedniej ilości zdjęć, albo wstawili za dużo zdjęć do jednej klasy, to zaowocowało tym że, klasy "piłka do tenisa" było o 3 za mało na szarej podłodze oraz o 2 za dużo na stole do ping-ponga. Ponadto zdjęć w klasie "piłka do siatkówki" na stole do ping-ponga jest o 1 za dużo, podobnie jeśli chodzi o zdjęcia na trawie, a zdjęć na podłodze jest o 1 za mało. Zdjęć do

klasy “piłka do ręcznej” jest o 1 za mało na stole do ping-ponga, a zdjęć do klasy “piłka do tenisa stołowego” jest o 1 za dużo na stole do ping-ponga

Przykłady zebranych przez nas danych:



Dodatkowo pozyskaliśmy 540 zdjęć ze zbiorów dostępnych w internecie. Zdjęcia są już opatrzone etykietą (każdemu zdjęciu przypisana jest dokładnie jedna klasa). Zebrane zdjęcia przedstawiają piłki z 6 klas uwzględnianych przez nasz system:

- piłka do nożnej - 90 zdjęć
- piłka do ręcznej - 90 zdjęć
- piłka do siatkówki - 90 zdjęć
- piłka do tenisa - 90 zdjęć
- piłka do koszykówki - 90 zdjęć
- piłeczka do tenisa stołowego - 90 zdjęć

Przykłady danych pozyskanych z internetu:



Podsumowując, zebraliśmy 540 zdjęć zebranych przez nas oraz 540 zdjęć znalezionych w internecie. Mamy więc do dyspozycji razem 1080 zdjęć.

#### **4. Zaplanuj procedurę adnotacji danych**

##### **1. Wybierz narzędzie do adnotacji, którego będziesz używać**

Będziemy używali narzędzia online do adnotacji Label Studio, ponieważ jest bezpłatne, intuicyjne w obsłudze i nie wymaga dużo czasu, aby nauczyć się z niego korzystać. Wydaje nam się, że to narzędzie sprosta naszym wymaganiom w ramach tego projektu.

##### **2. Stwórz precyzyjne instrukcje dotyczące wymagań związanych z adnotacją danych, sytuacji typowych i przypadków brzegowych**

1. Obejrzyj zdjęcie na kolorowym monitorze o średnicy co najmniej 13".
2. Zlokalizuj na zdjęciu obiekt będący piłką i określ klasę obiektu:
  - a. Jeżeli na zdjęciu znajdzie się jedna rozpoznawalna piłka, określ jej klasę
  - b. Jeżeli zdarzy się, że na zdjęciu nie będzie żadnych piłek, określ jej klasę jako "inne"
  - c. Jeżeli na zdjęciu znajdzie się piłka nieroznawalnego przez nas typu, określ jej klasę jako "inne"
  - d. Jeżeli zdarzy się, że na zdjęciu znajdzie się więcej niż jedna piłka, spójrz na tą znajdującą się bliżej kamery i przyporządkuj jej odpowiednią klasę
  - e. Jeżeli zdarzy się, że na zdjęciu znajduje się część piłki (niecała piłka), określ jej klasę
  - f. Jeżeli zdarzy się, że na zdjęciu jakiś obiekt, który zasłania część piłki, wybierz jej klasę jako "inne"
3. Określ czy na zdjęciu znajdują się inne obiekty poza piłką sportową

- a. Jeżeli poza rozpoznawalną piłką na zdjęciu znajdą się inne obiekty, zignoruj je i określ klasę piłki.
4. Określ jakość zdjęcia:
  - a. Jeżeli cień przysłania piłkę, nie sugeruj się nim i określ klasę piłki.
  - b. Jeżeli piłka jest niewyraźna lub w ruchu, przyporządkuj jej klasę.
5. Przypadki brzegowe
  - a. Jeżeli nie jesteś pewien do której z określonych wcześniej klas należy piłka, przypisz ją do klasy "difficult".
  - b. Jeżeli zmienisz zdanie, kliknij w odpowiednią etykietę klasy zgodną z nowym wyborem.
6. Jeżeli powyższe warunki są spełnione i odpowiednia klasa została przypisana do obiektu, kliknij 'Submit' przejdź do kolejnego zdjęcia poprzez kliknięcie na kolejny element z listy.
7. Jeżeli po zatwierdzeniu, stwierdzisz, że etykieta powinna być inna, wybierz odpowiednie zdjęcie z listy, zmień zaznaczoną etykietę i kliknij 'Update'.

### **3. Podziel surowe dane między członków zespołu, w celu ich adnotowania**

Członkowie zespołu:

- a. Aleksandra Rumińska - Zdjęcia, których właścicielem jest Przemysław Rośleń (określone w kolumnie "Owner").
- b. Przemysław Rośleń - Zdjęcia których właścicielem jest Norbert Nieżorawski (określone w kolumnie "Owner").
- c. Norbert Nieżorawski - Zdjęcia których właścicielem jest Aleksandra Rumińska (określone w kolumnie "Owner").

### **5. Dokonaj adnotacji na danych stosując zdefiniowaną procedurę. Przedstaw kilka przykładów i opisz problemy, jakie napotkałeś**

Przykładowe adnotacje zdjęć dokonane przez członków zespołu:



Etykieta: piłka do koszykówki



Etykieta: piłka do ręcznej



Etykieta: piłka do tenisa



Etykieta: piłka do nożnej



Etykieta: piłeczka do tenisa stołowego



Etykieta: piłka do siatkówki

Adnotowanie zdjęć nie obyło się bez problemów.

Napotkane problemy:

- Trudności w odróżnieniu piłek do nożnej od piłek do ręcznej, szczególnie jeżeli na zdjęciu nie było żadnych innych obiektów odniesienia, po których moglibyśmy ocenić ich rozmiar
- Na jednym zdjęciu widać tylko część piłki i dodatkowo zdjęcie jest słabej jakości, ponieważ telefon był w ruchu podczas robienia tego zdjęcia
- Niektóre zdjęcia były niskiej jakości przez co ciężej było rozpoznać piłkę na zdjęciu
- Niektóre zdjęcia piłek występują na niejednolitym tle (ręka, buty albo inne przedmioty)
- Niektóre piłki były zużyte, przez co trudno je było przydzielić do odpowiedniej etykiety
- Niektóre piłki były tak ułożone na zdjęciach, aby nie było widać ich cech charakterystycznych (np. łączeń, napisów, rozmiaru).
- Czasem występował bias, ze względu na to, że takie same piłki występowały w sklepach, w których zdjęcia były robione, przez co ich etykiety była anotatorowi z góry znana.
- Niektórzy użytkownicy mogli adnotować zdjęcia w określonej kolejności, zdefiniowanej przez nazwę zdjęcia (nie były one losowo dobierane), przez co mógł powstawać bias związany z oceną zdjęcia na podstawie wcześniejszego wzorca (seria piłek odpowiedniego typu)

Ostatecznie dokonaliśmy adnotacji 540 obrazów:

Klasa	Liczba adnotacji
piłka do nożnej	49
piłka do ręcznej	67
piłeczka do tenisa stołowego	90
piłka do koszykówki	81
piłka do siatkówki	74
piłka do tenisa	88
difficult	91
inne	0

## 6. Pomiar inter-observer variability/reliability

- a) W trakcie powyżej opisanej fazy adnotacji, 91 obrazów zostało sklasyfikowanych jako "difficult", czyli klasę przyznawaną w przypadku, gdy annotator nie był w stanie jednoznacznie stwierdzić, do jakiej klasy piłek sportowych zaliczyć obiekt na podanym zdjęciu. Przykładowe zdjęcia sklasyfikowane jako "difficult":



- b) Wszystkie 91 obrazów wspomnianych powyżej zostało przekazanych każdemu członkowi zespołu do indywidualnej adnotacji zgodnie z regułami przedstawionym w sekcji 4.2.
- c) Zgodność adnotacji: 86/91

Do adnotacji użyliśmy uprzednio wyekstrahowanych plików z narzędzia label-studio w formacie csv. Decyzję podejmowaliśmy wtedy kiedy etykieta była oznaczona "difficult". Oficjalna decyzja, została podejmowana, kiedy  $\frac{2}{3}$  głosów wskazywało na tą samą klasę.

Po ponownej adnotacji trudnych przykładów, oznaczonych etykietą "difficult" przez wszystkich członków zespołu uzyskaliśmy zgodność na poziomie około 94,5%. Oznacza to, że wystąpiło tylko kilka przypadków, a dokładnie 5, w których zdania były odmienne oraz którym przypisano różne etykiety. Najwięcej niezgodności pomiędzy klasami zaobserwowaliśmy przy klasach "piłka do nożnej" oraz "piłka do ręcznej", gdyż pojawiły się aż trzy takie obserwacje, co stanowiło 60% wszystkich niezgodnych obserwacji. Nie znalazł się żaden przykład, dla którego każdy annotator przypisał inną etykietę.

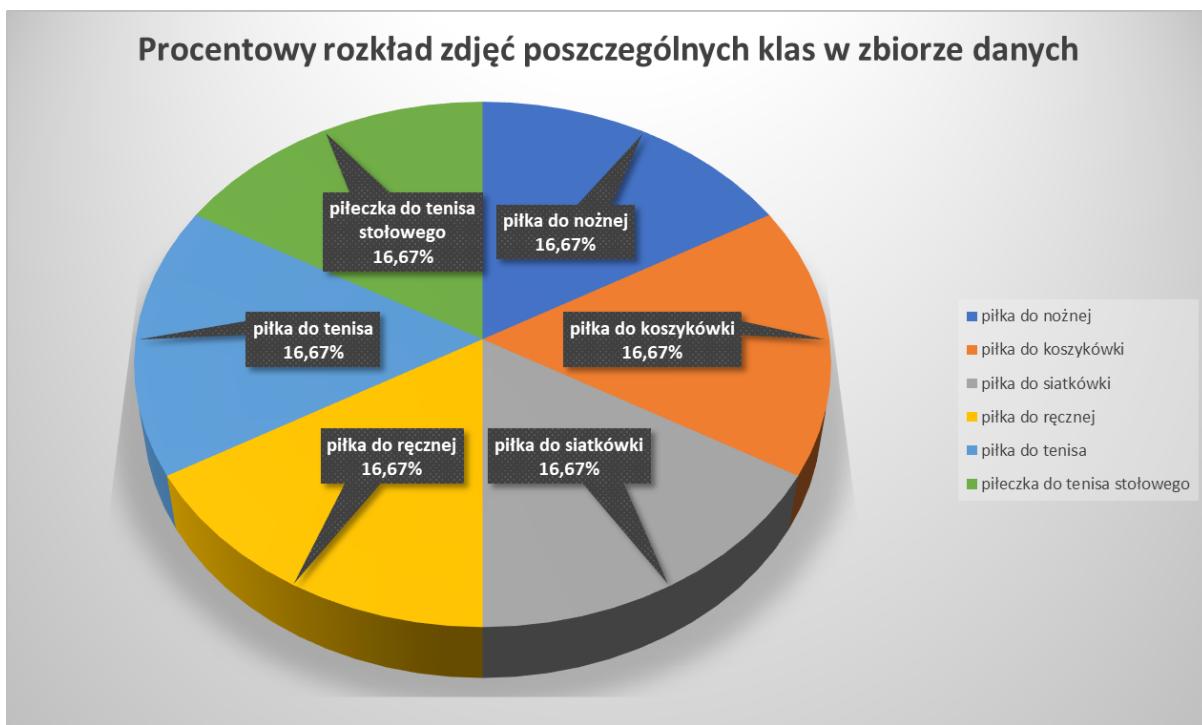
## 7. Przeanalizuj utworzony zbiór danych

- a) Statystyki

Cały zbiór danych zawiera 1080 zdjęć RGB (540 zdjęć zebranych przez nas oraz 540 pozyskanych z internetu), wśród których występuje zbalansowany podział na 6 klas: piłka do nożnej, piłka do ręcznej, piłka do siatkówki, piłka do koszykówki, piłeczka do tenisa stołowego, piłka do tenisa. Zatem, w zbiorze jest 180 zdjęć piłek z każdego rodzaju. Dodatkowo piłki na zdjęciach zrobionych przez nas przedstawione są na 3 różnych typach tła: podłożu betonowym, trawie/podłożu trawiastym i stole do tenisa stołowego. Każda z piłek ma dokładnie po 30 zdjęć na każdej z wymienionych powierzchni.

Warto zaznaczyć, że jest to zbiór danych z etykietami, które zostały im przydzielone przez twórców zbioru. W każdym wypadku tylko jedna etykieta została przypisana do jednego zdjęcia piłki. Wykres zaprezentowany poniżej opisuje liczbę zdjęć poszczególnych klas zarówno w zbiorze zdjęć zrobionych przez nas (540 zdjęć), jak i zbiorze zdjęć pozyskanych z internetu (540 zdjęć).





- b) Przeanalizuj relacje i korelacje w danych (grupy danych, podobne dane, korelacje między etykietami / klasami, źródła, itd.) Zaprezentuj kilka przykładów.

W każdej klasie możemy znaleźć podobne piłki na zdjęciach, po pierwsze ze względu na takie same 3 rodzaje otoczenia (w przypadku zdjęć robionych przez nas), które były w tle naszych zdjęć piłek, po drugie na umiejscowienie piłki na środku zdjęcia. Ponadto, duża ilość zdjęć ma mocno przybliżoną perspektywę, co sprawia, że piłki różnych rozmiarów mogą się wydawać takiej samej wielkości. Poniżej załączone zostały 2 przykładowe zdjęcia prezentujące kolejno piłkę do nożnej i do ręcznej, które w rzeczywistości znacznie różnią się rozmiarami, a na zdjęciach nie jest to zupełnie widoczne przez brak dodatkowych punktów odniesienia i podobną perspektywę.



Największy problem będą stanowiły dane należące do klas "piłka do nożnej" oraz "piłka do ręcznej". Oba rodzaje piłek są bardzo podobne do siebie jeśli chodzi o wykonanie, natomiast tym co pozwala nam je odróżnić w rzeczywistości jest ich rozmiar. Mając na uwadze wspomniany przez nas powyżej i widoczny na powyższych zdjęciach efekt, przewidujemy, że próbki należące do tych dwóch klas będą bardzo problematyczne dla modelu.

W przypadku zdjęć ze zbiorów internetowych większość zdjęć piłek do piłki ręcznej będzie na tle parkietu/sztywnego podłoża, a piłek do piłki nożnej na tle trawy, więc możliwe, że model będzie w stanie lepiej klasyfikować zdjęcia do kategorii, ale ze względu na przetrenowanie i dostosowanie się do tła, na którym dana piłka była ustawiona.

Zdjęcia piłek pozostałych klas wydają się być w miarę unikalne, ze względu na wykorzystany materiał, charakterystyczne wzory lub specyficzną dla danych rodzajów piłek kolorystykę.

Klasę "piłka do koszykówki" powinniśmy odróżnić po wzorach na piłce, gdyż są one bardzo unikalne. Równolegle zwykłe czarne linie, które łączą się ze sobą w dwóch miejscach. Ponadto, bardzo często piłki tego rodzaju są w kolorystyce różnych odcieni pomarańcza lub brązu. Nie jest tak jednak w każdym przypadku. Poniżej zaprezentowane są 3 przykładowe zdjęcia:



Jeśli chodzi o zdjęcia klasy "piłka do tenisa", są one odróżnialne ze względu na charakterystyczny materiał z małymi włoskami, z którego piłki zostały wykonane oraz charakterystyczną żółtawą barwę. Ponadto, piłki te w znacznej większości przypadków posiadają 2 równoległe, na ogół białe linie, które są ze sobą połączone. Dodatkowo, na tych piłach zazwyczaj umieszczony jest napis. Są to cechy, które naszym zdaniem zostaną wyłapane przez model uczący i pozwolą na względnie dobre odróżnienie próbki tej klasy od reszty. Poniżej zaprezentowane są 3 przykładowe zdjęcia:

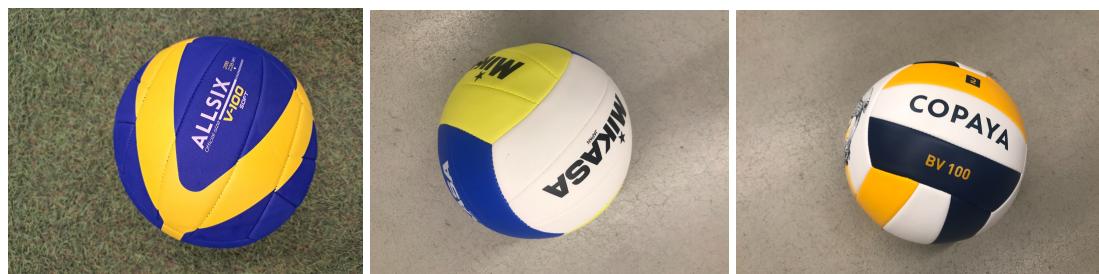


Klasę "piłeczka do tenisa stołowego" powinniśmy odróżnić przede wszystkim po rozmiarze. Na większości zdjęć, które zrobiliśmy, piłeczki te zajmują stosunkowo mniejszą przestrzeń na zdjęciu niż pozostałe co ma swoje odzwierciedlenie w rzeczywistości. Taka zależność w naszym zbiorze uczącym w zależności od sytuacji może okazać się zaletą lub wadą. Z jednej strony, model może nauczyć się, że piłeczki te są małych rozmiarów i będzie w stanie lepiej klasyfikować zdjęcia, na których piłeczki do tenisa stołowego będą zajmować względnie małą przestrzeń na zdjęciu, z drugiej strony, może mieć problem z poprawną klasyfikacją zdjęć, na których piłeczki te będą zajmować bardzo duży obszar zdjęcia. Kolejnymi cechami charakterystycznymi piłeczek do tenisa stołowego jest ich kolorystyka oraz materiał, z którego są wykonane. Są one w większości wykonane z plastiku i nie posiadają żadnych specjalnych wzorów, oprócz nazwy producenta. Jeżeli chodzi o

kolorystykę to są one w całości pokryte w jednolitym kolorze - w znacznej większości przypadków białym lub pomarańczowym. Są jednak nieliczne odstępstwa od tej reguły, np. na pierwszym z 3 zdjęć zaprezentowanych poniżej barwa przypomina raczej mało nasycony fiolet.



Ostatnią klasą piłek jest "piłka do siatkówki". Możemy ją odróżnić przede wszystkim po wzorze jakim są połączone ze sobą "prostokąty". Są one względem siebie równoległe lub prostopadłe, a oddzielone są poprzez nie rzucające się w oczy wcięcia o barwie pokrywającej się z barwą piłki w danym punkcie. Często na piłkach tego rodzaju dominują kolory: żółty oraz niebieski, jednak nie jest to reguła i nasz zbiór zawiera wiele piłek do siatkówki w innej kolorystyce. Ponadto, w wielu przypadkach jeden kolor jest dominujący na piłce, a drugi jest wykorzystany w charakterystycznych pasmach o nieregularnych kształtach tak jak na pierwszym ze zdjęć poniżej, jest też sporo przypadków gdy wspomniane wyżej "prostokąty" są zabarwione jednolicie, lecz mogą być realizowane w różnych kolorach w zależności od kolorystyki zastosowanej w danej piłce. To podejście widać wyraźnie na drugim oraz trzecim ze zdjęć poniżej.



Wszystkie zdjęcia każdej z klas piłek zostały przez nas wykonane na trzech rodzajach tła. Były to: betonowa nawierzchnia w kolorze szarym, powierzchnia stołu do tenisa stołowego o kolorze mało intensywnej czerni lub brązowawym oraz podłoże trawiaste o zielonym kolorze. 19 zdjęć tego ostatniego typu zostało wykonanych na dworze na tle naturalnej trawy, reszta z nich na tle sztucznej trawy. Znaczna większość zdjęć została wykonana w różnych sklepach sieci Decathlon. Postanowiliśmy robić zdjęcia na trzech różnych tła, aby ograniczyć w pewien sposób efekt polegający na nauczeniu się przez sieć neuronową konkretnego rodzaju tła zamiast lub oprócz piłki znajdującej się na zdjęciu. Dzięki pewnej różnorodności tła, istnieje większa szansa, że model będzie uczyć się głównie obiektów piłek na zdjęciach, bez większego uwzględniania tła.

Zdjęcia z internetu posiadają bardzo mocno zróżnicowane tła i scenerie. Nie ma tu wielkich ograniczeń co do nawierzchni dla danych piłek, jak to było w przypadku własnych zdjęć.

W zebranym przez nas zbiorze danych nie ma żadnych duplikatów. Jednakże niektóre zrobione przez nas zdjęcia przedstawiają tą samą instancję piłki na tym samym tle oraz posiadają podobną perspektywę przez co mogą sprawiać wrażenie, jakby były duplikatami. Takie zdjęcia nie są jednak dominującą grupą w naszym zbiorze, więc nie powinny stanowić większego problemu w procesie trenowania modelu.

W zbiorze zdjęć internetowych również nie znaleźliśmy żadnych duplikatów, chociaż bardzo często zdjęcia wyglądają podobnie, czasem nawet niemalże tak samo (ze względu brak tła albo gdy tło zajmuje zaniedbywalną część zdjęcia).

- c) Przeanalizuj zakłócenia i błędy w danych (rodzaje zakłóceń, rodzaje i liczba błędów, błędne etykiety, itd.) Zaprezentuj kilka przykładów.

Zakłócenia i błędy w danych zebranych przez nas:

- Najczęściej występującym rodzajem zakłóceń jakie występują zebranych przez nas danych są cienie oraz zaburzenia kolorów spowodowane nieodpowiednim oświetleniem w miejscach wykonywania zdjęć lub nieumiejemnym sposobem robienia zdjęć przez członków grupy projektowej.
- Na części zdjęć piłka nie znajduje się w centralnej części zdjęcia.
- Na części zdjęć piłka zajmuje stosunkowo mały obszar na zdjęciu (głównie dotyczy to zdjęć piłeczek do tenisa stołowego).
- Na niektórych zdjęciach poza piłką znajdują się również inne obiekty, które mogą zmylić sieć neuronową i skupić jej uwagę na sobie. Jednakże w niektórych przypadkach mogą one również stanowić różnorodność tła i przyczyniać model do tego, że zdjęcia które będą później wgrywane przez klientów systemu również mogą zawierać inne obiekty.
- Na niektórych zdjęciach odnotowaliśmy również zmniejszenie ostrości spowodowane np. zbyt bliską lub zbyt daleką odlegością piłki od aparatu.
- Na niewielu zdjęciach wystąpiło również rozmycie spowodowane poruszeniem aparatu podczas robienia zdjęcia.
- Na jednym zdjęciu widoczna jest tylko część piłki.

Podczas etykietowania zebranych przez nas danych znalazło się 91 przypadków trudnych do rozstrzygnięcia. Etykiety zostały nadane tym zdjęciom poprzez głosowanie członków grupy projektowej. Nadawana została etykieta, która otrzymała więcej głosów. Po adnotacji danych i porównaniu przydzielonych zdjęciom etykiet ze stanem faktycznym, okazało się, że wszystkie etykiety zostały przypisane poprawnie.

Zakłócenia i błędy w danych pozyskanych z internetu:

- Najczęściej występującym rodzajem zakłóceń jaki występuje w danych pozyskanych z internetu są zniekształcenia geometryczne piłek na zdjęciach spowodowane przeskalowaniem tych zdjęć do rozdzielczości 224x224
- Kolejnym często występującym rodzajem zakłóceń jaki odnotowaliśmy jest mała ostrość wielu zdjęć.

- Na wielu zdjęciach poza piłką znajdują się również inne obiekty, które mogą zmylić sieć neuronową i skupić jej uwagę na sobie. Jednakże w niektórych przypadkach mogą one również stanowić różnorodność tła i przyzwyczać model do tego, że zdjęcia które będą później wgrywane przez klientów systemu również mogą zawierać inne obiekty.
- Na części zdjęć piłki stanowiły mały procent całego zdjęcia oraz nie znajdowały się w centralnej jego części
- Na części zdjęć widoczna jest tylko część piłki

- d) Zbadaj, czy niektóre dane są "trudniejsze" lub "ważniejsze" od innych, zdecyduj, czy trzeba je traktować oddzielnie (np. nadawanie wag podczas treningu lub oddzielne zbiory testowe).

Większość obrazów znajdujących się w zebranym zbiorze jest wyraźna, dobrej jakości i można jednoznacznie stwierdzić ich klasę. Znajdują się jednak w zbiorze danych zdjęcia, które posiadają niektóre z wymienionych powyżej zakłóceń przez co będą trudniejsze dla sieci neuronowej w procesie trenowania modelu. Na szczęście liczba zdjęć obarczonych zakłóceniami jest stosunkowo niewielka oraz poziom występujących zakłóceń nie jest na tyle duży, aby w znaczący sposób zakłócić proces treningu modelu. Zdecydowaliśmy się więc nie wykluczać tych zdjęć ze zbioru treningowego.

## 1. Zbliżenia

Niektóre zdjęcia w zbiorze danych przedstawiają piłki ze zbliżeniem 2x albo większym, a są też takie, które zostały zrobione przy użyciu oddalenia 0.5x. Przez zastosowanie tych ustawień piłka do tenisa stołowego, z natury bardzo niewielka, mogła wyglądać, jakby była takiego samego rozmiaru jak na przykład piłka do siatkówki. W drugą zaś stronę, wykorzystanie oddalenia pozwoliło sprawić, że piłka do piłki nożnej wydawała się mniejsza i mogłaby zostać pomylona z piłką do ręcznej. Były to celowo stosowane zabiegi, aby sama wielkość piłki nie była powodem, dla którego obraz zostałby sklasyfikowany przez model do danej klasy. Dlatego też zdjęcia z różnymi przybliżeniami nie będą traktowane inaczej od pozostałych zdjęć ze zbioru.

## 2. Nieostrość i piłki w ruchu

W zbiorze danych zebranych przez nas występuje kilka zdjęć, które są nieostre, ponieważ piłka jest w ruchu. Powoduje to, że nie są widoczne wszystkie szczegóły jej powierzchni, takie jak dokładny kształt szwów, napisy czy materiał, z którego została wykonana. Natomiast to poruszenie nie zaburza kształtu piłki, jej wielkości, czy przybliżonego rozmieszczenia kolorów. Jest to na tyle wystarczające do identyfikacji piłki, że w fazie adnotacji nie pojawiły się problemy z przydzieleniem tym zdjęciom niewłaściwych etykiet. Zatem nie są to zdjęcia na tyle trudne, żeby wymagały specjalnego traktowania.

W zbiorze danych pozyskanych z internetu nie braliśmy pod uwagę piłek, które byłyby w ruchu, więc problem piłki w ruchu nie występuje. Natomiast dość dużo zdjęć, musiało zostać przeskalowanych w górę, przez co straciły one na ostrości.

### 3. Niepełne piłki

W zbiorze danych zebranych przez nas znalazło się również dokładnie jedno zdjęcie, na którym piłka była po pierwsze poruszona, co zostało omówione powyżej, jak również tylko jej połowa znajduje się na zdjęciu. Jednak z powodu istnienia tylko jednego takiego przypadku, zdecydowaliśmy się uwzględnić go w zbiorze danych, żeby był on zbilansowany i zawierał równą ilość zdjęć z każdej klasy.

W zbiorze danych pozyskanych z internetu znajduje się część piłek które nie są widoczne w pełni. Przypadek ten występuje, gdy piłka ręczna jest trzymana w ręku, więc część ciała zasłania kawałek obwodu piłki. Ponadto, na niektórych zdjęciach piłka znajduje się w siatce, którą ją częściowo zakrywa.

### 4. Inne przedmioty na zdjęciu

Kolejny problem, który został wykryty w zbiorze danych zebranych przez nas to obecność innych obiektów na zdjęciach, takich jak bramki, fragment nogi czy ręki. W żadnym jednak przypadku nie jest to fragment innej piłki tego samego lub różnego typu. W dodatku przedmioty te nie zajmują większości obszaru zdjęcia, a jeśli zakrywają jakiś fragment piłki to w nieznacznym stopniu, tak, że może ona nadal zostać rozpoznana bez problemu. Nie będziemy zatem stosować specjalnych niższych wag, aby zmniejszyć istotność tych zdjęć.

W zbiorze danych pozyskanych z internetu znajdują się zdjęcia, na których są obecne inne przedmioty poza piłką. Przykładem tego jest rakietka do tenisa stołowego (albo jej kawałek) widoczna na zdjęciu razem z piłeczką. Nie jest to szczególnie problem, gdyż przyszli użytkownicy takiego systemu często mogliby wysyłać zdjęcia, na których poza piłką widoczny byłby inny przedmiot ściśle powiązany z danym sportem. Dodatkowo, są również pojedyncze zdjęcia, na których widoczne są części innych piłek danego rodzaju. W takich przypadkach jedna piłka jest widoczna w całości i zajmuje znakomitą część obrazka. Ostatecznie zdecydowaliśmy się nie uwzględniać tych zdjęć w procesie uczenia modelu.

### 5. Różnorodna nawierzchnia trawiasta

Około 33% piłek z każdej klasy, a zatem 33% wszystkich piłek w zbiorze danych zebranych przez nas zostało zrobionych na "trawie". Jednakże, w skład tej kategorii wchodziły zarówno żywa trawa rosnąca na dworze, jak również sztuczna nawierzchnia w sklepie, która miała ją imitować. Co więcej, zdjęcia wykonane na nawierzchni w sklepie przy użyciu różnego przybliżenia miały różne odcienie zieleni - od szarozielonego po zielononiebieski. Uznaliśmy, że mimo tych różnic w kolorach nie będziemy specjalnie traktować tych zdjęć, gdyż przy innych tła kolory również wychodziły mniej lub bardziej przyjmione ze względu na oświetlenie. Różnorodne kolory tła na zdjęciach mogą pozytywnie wpływać na proces treningu, minimalizując efekt uczenia się przez sieć konkretnych rodzajów tła co nie jest przez nas.

### 6. Format zdjęć

Ostatnim problemem wpływającym na ocenę zdjęć, jest format zdjęć, gdyż większość z nich została zrobiona w takim samym formacie 4000x3000, Natomiast około 6% zdjęć ma format 4032 x 1908. Jako że w tym formacie grupy klas nie są równo rozłożone, klasyfikator widząc taz piłką w podobnym formacie, może już przypisywać automatycznie wyuczoną klasę. Trzeba to przekształcić na format 4000 x 3000 aby takie taka sytuacja nie miała miejsca.

W zbiorze danych pozyskanych z internetu, część zdjęć miała już rozdzielcość 224x224. Natomiast była to tylko niewielka część zdjęć, która nadawała się do wykorzystania w modelu bez zmiany rozdzielcości. Większość zdjęć posiadała bardzo różne rozdzielcości, czasem większe niż 224x224, a czasem mniejsze.

#### 7. Piłki stanowiące niewielką część zdjęcia i nie będące w centralnej jego części

W zbiorze danych pozyskanych z internetu znajdowało się wiele zdjęć, na których piłki stanowiły niewielką część całości zdjęcia i/lub nie były w umiejscowione w centralnej części zdjęcia. Z tego względu wykonane zostało przycięcie zdjęcia, aby piłka stanowiła znakomitą część docelowego zdjęcia oraz była w centrum.

Ostatecznie ustaliliśmy, że zdjęcia o innym formacie wymagają pewnych przekształceń, jak zostało to opisane powyżej, po zastosowaniu których staną się równie ważne jak pozostałe przykłady znajdujące się w zbiorze danych. Zatem cały zbiór stanie się jednakowo ważny i nie będzie w nim przypadków, które będą wymagały innego specjalnego traktowania, czy dostosowania wag.

## 8. Przetwarzanie danych

### a. Wybierz metodę reprezentacji danych. Skonwertuj dane.

Dane reprezentowane będą jako trójwymiarowe tensory, o wymiarach (wysokość, szerokość, liczba kanałów). Obrazy będą w formacie RGB, a zatem liczba kanałów będzie wynosiła 3, po jednym dla każdego z kolorów: czerwony, zielony, niebieski. Biorąc pod uwagę, że modele do przetwarzania obrazów często wymagają, aby miały one kształt kwadratu, wartości wysokość i szerokość będą sobie równe. Zdecydowaliśmy się użyć rozdzielcości 224x224, bardzo popularnej w dzisiejszych modelach opartych na sieciach neuronowych. Zatem uzyskane przez nas tensory będą miały kształt: (1080 + 1080\*n, 224, 224, 3). Pierwszy człon, równy 1080 + 1080\*n, odpowiada sumie liczby oryginalnych zdjęć oraz liczbnych zdjęć, które zostaną wygenerowane przy pomocy różnych funkcji modyfikujących nasze bazowe dane (proces augmentacji). Zmienna 'n' określa liczbę zastosowanych funkcji augmentacji danych. Uzyskane dane zostaną połączone wraz z ich etykietami do obiektu `tf.data.Dataset` przy pomocy funkcji `from_tensor_slices()`.

Procedura dla zdjęć wykonanych przez nas:

1. Jako pierwsze, na podstawie arkusza kalkulacyjnego z wynikami adnotacji, gdzie do nazwy zdjęcia przypisana była etykieta, utworzyliśmy foldery o nazwach takich samych jak klasy i przydzieliśmy do nich zdjęcia. Następnie na podstawie przeglądu tła, rozdzieliśmy zdjęcia z każdego katalogu na trzy oddzielne, każdy z nich odpowiadał innej nawierzchni. W ostateczności mieliśmy 18 katalogów, po 3 dla

każdej z klas. Taki podział na katalogi, uwzględniający tło ułatwi podział danych na zbiory treningowy, walidacyjny i testowy, gdyż chcemy, aby w każdym z nich zbalansowana była liczba piłek względem zarówno ich klas, jak i tła.

2. Przy pomocy funkcji `os.walk()`, pobieramy wszystkie katalogi, a następnie iterując po nich, wybieramy ich zawartość z rozszerzeniem `.jpg`, i dodajemy do poszczególnych zbiorów, jako etykietę zapisując nazwę folderu.

```
# pobranie folderów=klas z folderu piłki
directories_classes = []

for root, dirs, _ in os.walk("./pilki", topdown=False):
    for name in dirs:
        directories_classes.append(name)
```

3. Po przejściu wszystkich folderów, dla każdego ze zbiorów wywołujemy funkcję, która ujednolici nazwy klas, usuwając z nich człon dotyczący nawierzchni, także wszystkie piłki jednej kategorii, na przykład piłki do gry w piłkę nożną, będą miały jedną i tą samą kategorię o nazwie “`pilka_do_noznej`”, niezależnie od tła, na którym się znajdują.

```
# Funkcja aby ujednolicić etykiety na koniec, pozbywając się członu trawa,
stol, podloga
def unify_labels(labels_array):
    for i in range(len(labels_array)):
        label_subs = labels_array[i].split('_')
        labels_array[i] = '_'.join(label_subs[:-1])
```

4. Na sam koniec tworzone są tensory, za pomocą których przedstawione są zdjęcia, oraz ich etykiety łączone są w jeden obiekt wspólny poprzez wykorzystanie funkcji: `tf.data.Dataset.from_tensor_slices()`.

```
# stwórz datasety train, val, test
dataset_train = tf.data.Dataset.from_tensor_slices((train_images,
train_labels))
dataset_validation = tf.data.Dataset.from_tensor_slices((val_images,
val_labels))
dataset_test = tf.data.Dataset.from_tensor_slices((test_images,
test_labels))
```

Procedura dla zdjęć pozyskanych z internetu:

1. Jako pierwsze, na podstawie arkusza kalkulacyjnego z wynikami adnotacji, gdzie do nazwy zdjęcia przypisana była etykieta, utworzyliśmy foldery o nazwach takich samych jak klasy i przydzieliliśmy do nich zdjęcia. W ostateczności powstało 6 katalogów, po 1 dla każdej z klas. Taki podział na katalogi, ułatwi podział danych na zbiory treningowy, walidacyjny i testowy, gdyż chcemy, aby w każdym z nich zbalansowana była liczba zdjęć względem ich klasy.
2. Ta sama procedura jak w punkcie 2. w przypadku zdjęć wykonanych przez nas.
3. Ta sama procedura jak w punkcie 4. w przypadku zdjęć wykonanych przez nas.

**b. Zaproponuj i wykonaj standaryzację/normalizację danych oraz inne niezbędne przetwarzanie wstępne**

Zdecydowaliśmy się zastosować normalizację, skalując wartość pikseli tak, by mieściły się w zakresie od 0 do 1, gdyż może ona pomóc w kwestiach takich jak redukcja efektów oświetlenia czy kontrastu. Należy zaznaczyć, że dane były w formacie RGB, a zatem znormalizowany zostanie każdy z kanałów.

Zebrane dane zostaną znormalizowane przy pomocy kodu poniżej:

```
def normalize_data(img, label):
    return tf.cast(img, tf.float32)/255., label

dataset_train = dataset_train.map(normalize_data)
dataset_validation = dataset_validation.map(normalize_data)
dataset_test = dataset_test.map(normalize_data)
```

**c. Zaproponuj i wykonaj rozszerzenie danych (augmentację danych)**

Przed dokonaniem augmentacji danych, aby zachować statystyki opisane w punkcie 7, musieliśmy usunąć 2 zdjęcia klasy "piłka do tenisa" (20230411\_131013.jpg; 20230411\_131107.jpg) na stole do ping-ponga, 1 zdjęcie klasy "siatkówka" (IMG\_1005.jpg) na stole do ping-ponga, 1 zdjęcie klasy "siatkówka" (20230411\_135919.jpg) na trawie oraz 1 zdjęcie klasy "piłeczka do tenisa stołowego" (20230411\_130605.jpg) na stole do ping-ponga.

Ponadto musieliśmy zduplikować 3 zdjęcia klasy "piłka do tenisa" (20230411\_125801.jpg; 20230411\_125644.jpg; 20230411\_12637.jpg) na powierzchni betonowej, 1 zdjęcie klasy "piłka do siatkówki" (20230411\_123039.jpg) na powierzchni betonowej oraz 1 zdjęcie klasy "piłka do ręcznej" (IMG\_1136.jpg) na stole do ping-ponga.

Zastosowaliśmy cztery sposoby augmentacji danych:

- flip - tworząc odbicie lustrzane obrazu względem osi x lub y, my zastosowaliśmy odbicie względem osi x

```
def flip_image(img, direction):  
    return cv2.flip(img, direction)
```

- zmiana jasności obrazu, my zastosowaliśmy przyciemnienie, delta = -0.4

```
def adjust_brightness(img, delta):  
    return tf.image.adjust_brightness(img, delta=delta)
```

- rotate - tworząc obraz obrócony o losową wartość kąta z podanego zakresu, my zastosowaliśmy zakres od -90 do 90 stopni

```
def rotate_image(img, angle_range_left, angle_range_right):  
    angle = np.random.randint(angle_range_left, angle_range_right)  
    rows, cols, _ = img.shape  
    M = cv2.getRotationMatrix2D((cols / 2, rows / 2), angle, 1)  
    return cv2.warpAffine(img, M, (cols, rows))
```

- zoom - tworząc obraz z efektem przybliżenia w zależności od parametru zoom\_factor, my zastosowaliśmy zoom\_factor = 2.2

```
def zoom_image(img, zoom_factor):  
    h, w = img.shape[:2]  
    new_h, new_w = int(h / zoom_factor), int(w / zoom_factor)  
    center_x, center_y = int(w / 2), int(h / 2)  
    cropped_img = img[center_y - new_h:center_y + new_h, center_x -  
new_w:center_x + new_w]  
    return cv2.resize(cropped_img, (w, h))
```

Poniższy kod prezentuje wykonanie przez nas tych operacji na danych oryginalnych zdjęciach i dodanie wygenerowanych obrazów do wspólnej listy z bazowymi.

```
#augemntacja danych - plus 4x tyle zdjęć  
#flip pionowy względem osi x  
flipped_img = flip_image(img, 0)  
#zmiana jasności (przyciemnienie)  
bright_img = adjust_brightness(img, -0.4)  
#obrót obrazu o losowy kąt z zakresu (-90, 90) stopni  
rotated_img = rotate_image(img, -90, 90)  
#zoom x 2  
zoomed_img = zoom_image(img, 2.2)  
  
# dodaj nowo powstałe dane do tablic  
image_array.append(flipped_img)
```

```
label_array.append(dir)
image_array.append(bright_img)
label_array.append(dir)
image_array.append(rotated_img)
label_array.append(dir)
image_array.append(zoomed_img)
label_array.append(dir)
```

Jako że każda z metod działa na wszystkich danych, niezależnie od klasy czy rodzaju tła, to ostatecznie otrzymaliśmy 5x więcej obrazów niż mieliśmy oryginalnie, co przekłada się na 5400 obrazów.

## 9. Podziel dane na zbiory: treningowy (TRAIN), walidacyjny (VAL) oraz testowy (TEST). Przedstaw kilka statystyk dla wygenerowanych zbiorów.

Mając na uwadze, że zbiór treningowy powinien być znacznie większy niż zbiór walidacyjny i testowy, aby model miał wystarczającą ilość danych, aby nauczył się rozpoznawać klasę piłek sportowych, podzieliliśmy dane w proporcji 80:10:10. Oznacza to, że 80% wszystkich danych zostało przydzielonych do zbioru treningowego, a po 10% do zbiorów walidacyjnego i testowego. Powyższe dotyczy splitu 1 i splitu 2. W splicie 3 dane zostały podzielone w stosunku 90% do zbioru treningowego, a 10% do zbioru testowego.

Ważnym aspektem przy podziale danych było także zapewnienie, że w każdym ze zbiorów wystąpią tensory opisujące zdjęcia piłek z każdej klasy i na każdym tle. Zostało to zapewnione w trakcie pobierania zdjęć z odpowiednich folderów, jak opisano w punkcie 8a. Przy każdej iteracji pętli po liście z rozszerzonymi o tło nazwami klas, dwa razy (jeden raz dla zbioru walidacyjnego i jeden dla testowego) wywoływana zostawała funkcja przedstawiona poniżej:

```
# Funkcja do tworzenia setu walidacyjnego i testowego
def build_subset(image_array, label_array, val_images, val_labels, size):
    random_indices = random.sample(range(len(image_array)), size)

    for i in random_indices:
        val_images.append(image_array[i])
        val_labels.append(label_array[i])

        image_array_removed = [image_array[index] for index in
range(len(image_array)) if index not in random_indices]

        label_array = np.array(label_array)
        label_array = np.delete(label_array, random_indices)
        label_array = label_array.tolist()
```

```

    return image_array_removed, label_array, val_images, val_labels

# wywołaj funkcję losowo przydzielającą do val i test setów
    image_array_split1, label_array_split1, test_images_split1,
test_labels_split1 = build_subset(image_array_split1,
                                label_array_split1,
                                test_images_split1,
                                test_labels_split1,
                                size_split1)

image_array_split1, label_array_split1, val_images_split1,
val_labels_split1 = build_subset(image_array_split1,
                                label_array_split1,
                                val_images_split1,
                                val_labels_split1,
                                size_split1)

image_array_split2, label_array_split2, test_images_split2,
test_labels_split2 = build_subset(image_array_split2,
                                label_array_split2,
                                test_images_split2,
                                test_labels_split2,
                                size_split2_3)
image_array_split2, label_array_split2, val_images_split2,
val_labels_split2 = build_subset(image_array_split2,
                                label_array_split2,
                                val_images_split2,
                                val_labels_split2,
                                size_split2_3)

image_array_split3, label_array_split3, test_images_split3,
test_labels_split3 = build_subset(image_array_split3,
                                label_array_split3,
                                test_images_split3,
                                test_labels_split3,
                                size_split2_3)

```

Przyjmowała ona jako argumenty listy wszystkich obrazów z danego folderu, ich etykiety, listy obrazów i etykiet odpowiednio dla zbioru walidacyjnego (dla splitów 1 i 2) lub testowego oraz liczbę, która odpowiadała 10% danych z klasy. To właśnie tyle losowych indeksów było pobieranych z listy posiadającej wszystkie obrazy z danej klasy. Następnie dane, które znajdowały się pod tymi indeksami były przepisywane do zbioru walidacyjnego (dla splitów 1 i 2) lub testowego i usuwane ze zbioru oryginalnego. Dane, które pozostawały w pierwotnym zbiorze, zostały dodane do zbioru treningowego - zostało to zaprezentowane poniżej.

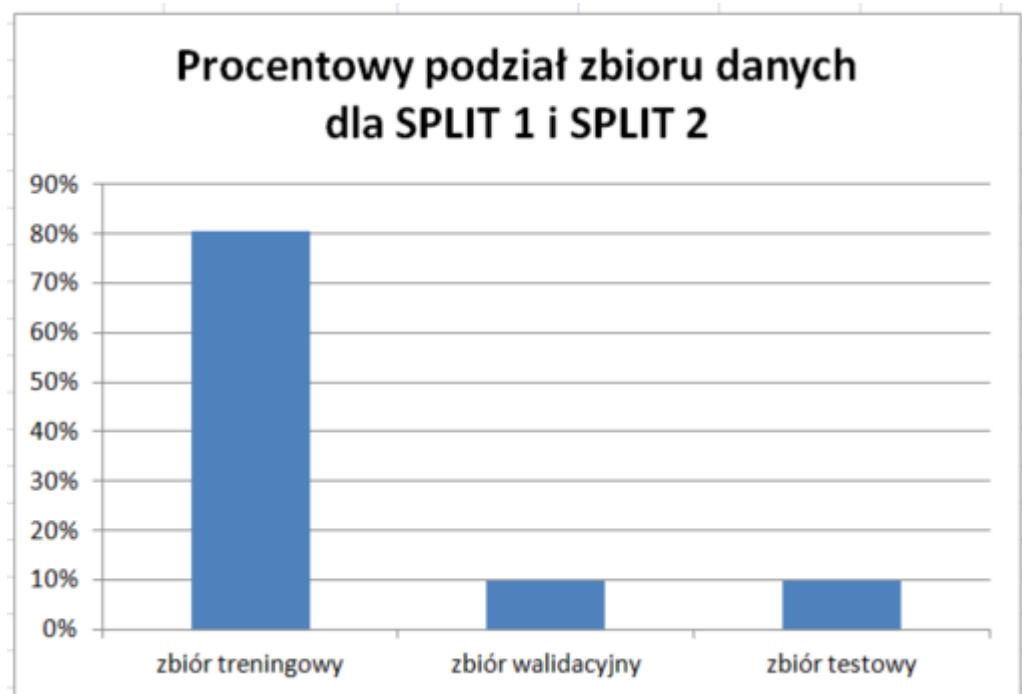
```
# rozszerz listę train o dane z tymczasowych tablic po usunięciu elementów
# przydzielonych do val i test

train_images_split1 = train_images_split1 + image_array_split1
train_labels_split1 = train_labels_split1 + label_array_split1

train_images_split2 = train_images_split2 + image_array_split2
train_labels_split2 = train_labels_split2 + label_array_split2

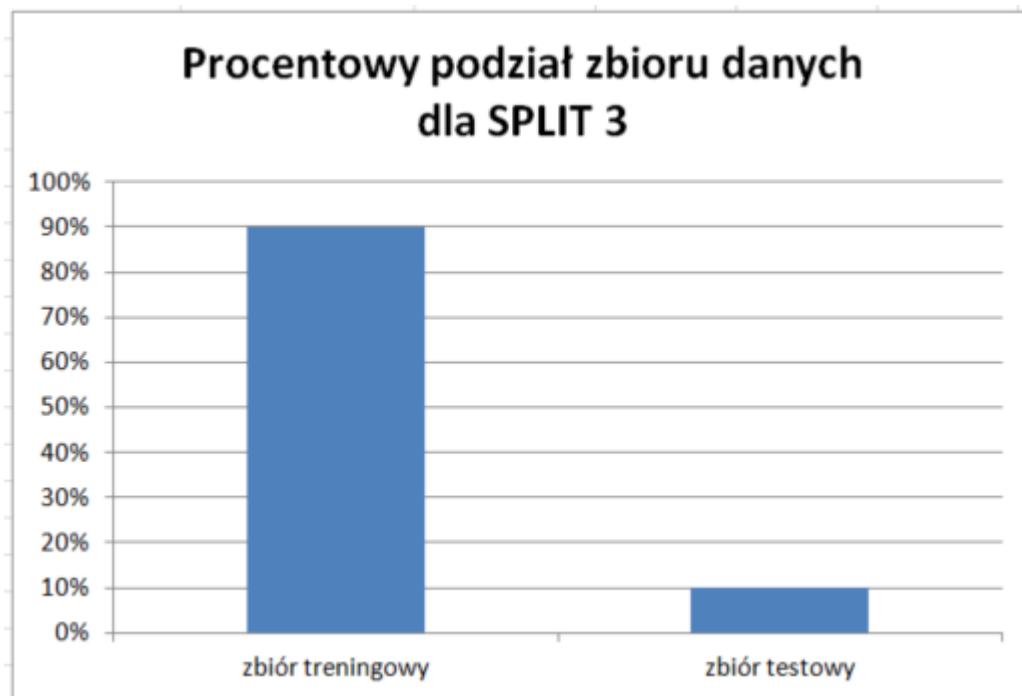
train_images_split3 = train_images_split3 + image_array_split3
train_labels_split3 = train_labels_split3 + label_array_split3
```

Poniższy wykres prezentuje procentowy podział zbioru danych na zbiory treningowe, walidacyjne i testowe dla splitów 1 i 2.



Jak widać w zbiorze treningowym jest 80% danych, w walidacyjnym 10% i testowym też 10%. Warto nadmienić, że w każdym z nich mamy równomierny rozkład reprezentantów różnych klas oraz równomierny podział zdjęć wygenerowanych w procesie augmentacji.

Poniższy wykres prezentuje procentowy podział zbioru danych na zbiory treningowe i testowe dla splitu 3.



Jak widać w zbiorze treningowym jest 90% danych, a w testowym też 10%, ponieważ zbiór testowy został wchłonięty w zbiór treningowy. Warto nadmienić, że w każdym z nich mamy równomierny rozkład reprezentantów różnych klas oraz równomierny podział zdjęć wygenerowanych w procesie augmentacji.