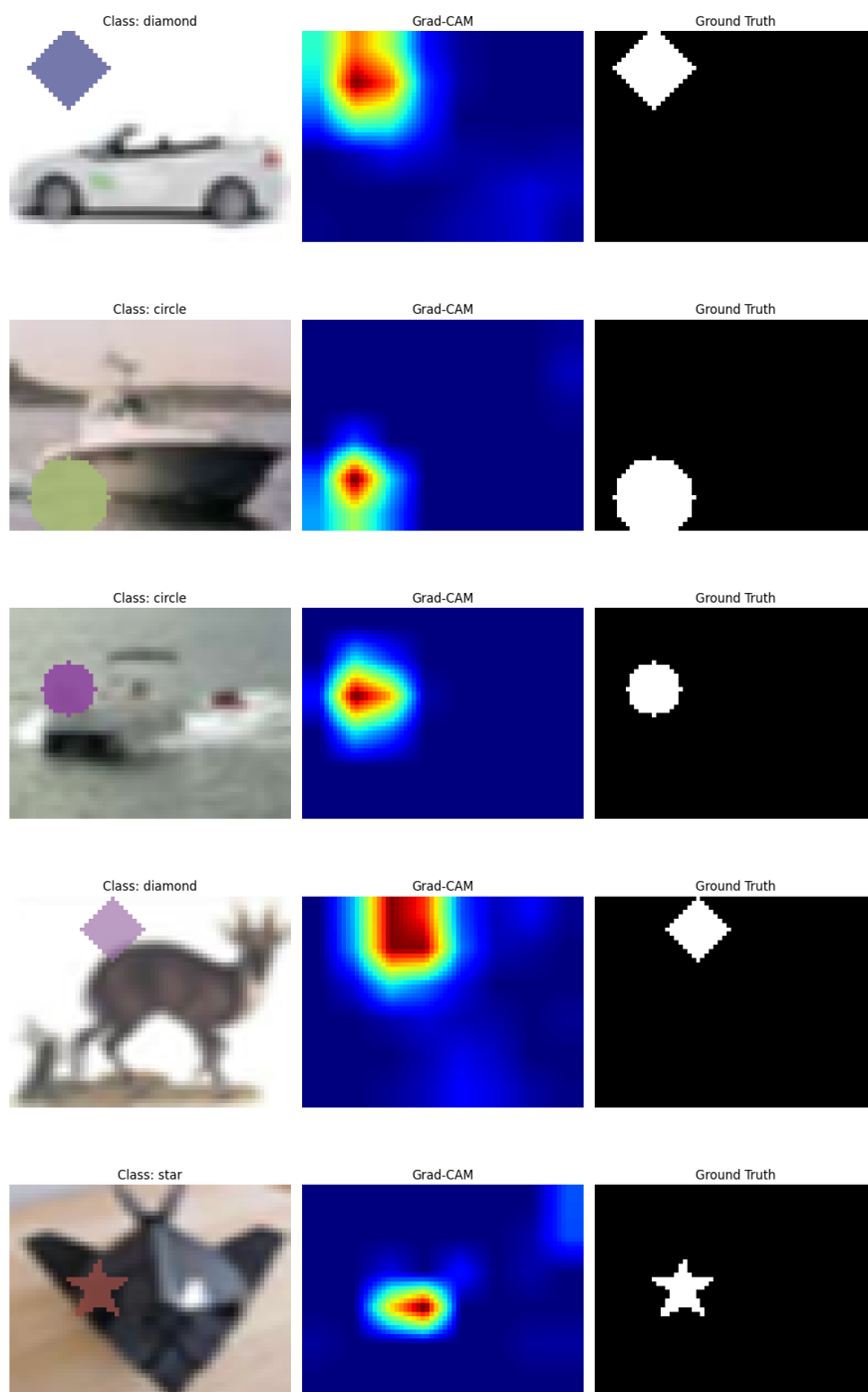


Raport: Segmentacja obrazów bez danych wejściowych użytkownika

Marysia Nazarczuk

13 stycznia 2026

1. Wizualizacja wyników Grad-CAM



Rysunek 1: Przykłady działania Grad-CAM dla różnych klas kształtów. Kolory od niebieskiego (niska istotność) do czerwonego (wysoka istotność) wskazują regiony najbardziej istotne dla decyzji modelu klasyfikującego.

Na Rysunku 1 przedstawiono działanie zaimplementowanej metody Grad-CAM. Jak można zaobserwować, heatmapy generowane przez algorytm skutecznie lokalizują kształty na obrazach, koncentrując się na regionach zawierających obiekty docelowe (koło, kwadrat, trójkąt, gwiazda, diament). Maksima aktywacji w heatmapach odpowiada pozycjom obiektów, co potwierdza poprawność implementacji.

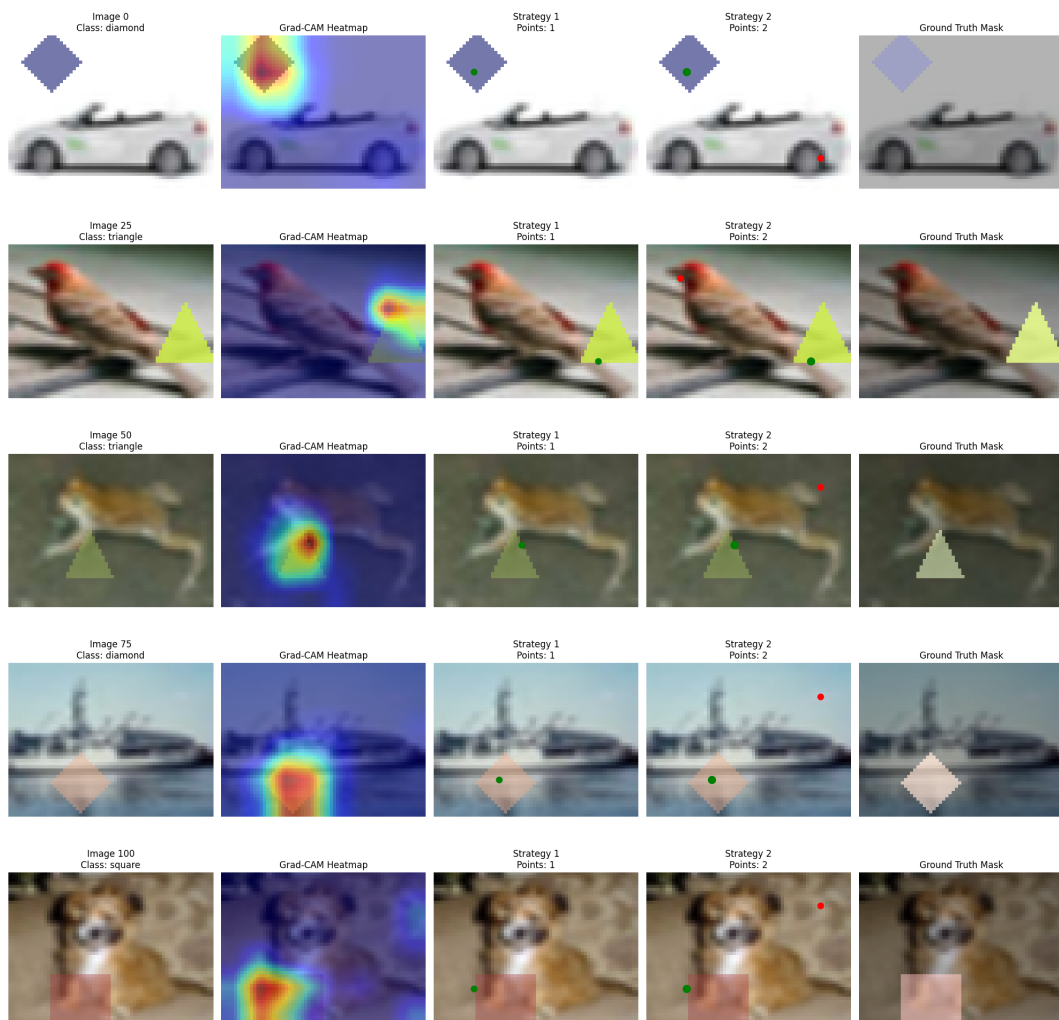
2. Opis strategii dla SAM

2.1 Strategia 1: Tylko punkty foreground

W pierwszej strategii wykorzystano heatmapę z Grad-CAM do automatycznego znalezienia punktów wskazujących na obiekt. Dla każdego obrazu wybierany jest punkt o maksymalnej wartości aktywacji w heatmapie, który z wysokim prawdopodobieństwem znajduje się wewnątrz docelowego kształtu. Punkt ten jest przekazywany do modelu SAM jako jedyna wskazówka typu foreground.

2.2 Strategia 2: Punkty foreground i background

Druga strategia rozszerza pierwszą o dodatkowe punkty background. Oprócz punktu z maksymalnej aktywacji z Grad-CAM (foreground), dodawane są punkty w rogach obrazu jako tło. Wybierany jest róg najbardziej oddalony od punktu foreground, co dostarcza modelowi SAM informacji o tym, które regiony obrazu z pewnością *nie* należą do obiektu.



Rysunek 2: Porównanie obu strategii na przykładowych obrazach: (a) oryginalny obraz, (b) heatmapa Grad-CAM, (c) strategia 1 (tylko foreground), (d) strategia 2 (foreground + background), (e) maska ground-truth. Zielone punkty oznaczają foreground, czerwone - background.

3. Metryki ewaluacyjne

3.1 Jakość generowanych punktów

Metryka	Strategia 1	Strategia 2
Hit Rate	68.4%	34.3%
Średnia odległość od środka maski	5.99 px	23.67 px

Tabela 1: Metryki jakości automatycznie generowanych punktów

Tabela 1 przedstawia jakość punktów generowanych przez obie strategie. Hit Rate (trafność) mierzy, jak często generowane punkty trafiają w obszar maski ground-truth. Strategia 1 osiąga trafność na poziomie 68.4%, co oznacza, że w około 2/3 przypadków punkt wybrany przez Grad-CAM faktycznie znajduje się wewnątrz obiektu. Strategia 2 ma znacznie niższą trafność (34.3%), ponieważ zawiera punkty background celowo umieszczone poza obiektem. Strategia 2 ma większą średnią odległość od środka maski (23.67 px vs 5.99 px), co wynika z uwzględnienia odległych punktów background.

3.2 Jakość segmentacji (IoU)

Strategia	Mean IoU	Std IoU	Min/Max IoU
Tylko foreground	72.6%	0.435	0.0 / 1.0
Foreground + background	72.1%	0.438	0.0 / 1.0

Tabela 2: Metryki Intersection over Union (IoU) dla obu strategii

W Tabeli 2 przedstawiono wyniki segmentacji mierzone metryką Intersection over Union. Obie strategie osiągają bardzo podobne wyniki na poziomie około 72% IoU, co znacząco przekracza wymagane 65%. Wysokie odchylenie standardowe (0.435 i 0.438) wskazuje na dużą zmienność wyników między różnymi obrazami. Podczas gdy dla większości obrazów segmentacja jest doskonała ($\text{IoU} = 1.0$), istnieją przypadki gdzie algorytm całkowicie nie radzi sobie z segmentacją ($\text{IoU} = 0.0$). Interesującym spostrzeżeniem jest fakt, że dodanie punktów background nie poprawiło jakości segmentacji, a nawet nieznacznie ją obniżyło (z 72.6% do 72.1%).

4. Dyskusja i możliwe ulepszenia

Zaproponowane rozwiązanie demonstruje skuteczne połączenie metody explainable AI (Grad-CAM) z zaawansowanym modelem segmentacji (SAM), umożliwiając w pełni automatyczną segmentację obiektów bez interwencji użytkownika.

Kluczowe obserwacje: 1. Grad-CAM lokalizuje obiekty z trafnością 68.4%, co jest zadowalającym wynikiem biorąc pod uwagę prostotę metody (wybór punktu o maksymalnej aktywacji). 2. Mimo że strategia 2 ma niższą trafność punktów (34.3%), jakość segmentacji (IoU) jest porównywalna ze strategią 1. Sugeruje to, że model SAM jest odporny na obecność niektórych nieprawidłowych punktów background. 3. Dodanie punktów background nie poprawiło jakości segmentacji w tym konkretnym zadaniu, co może wynikać z faktu, że obiekty są względnie izolowane na jednolitym tle. 4. Główne źródło błędów stanowią przypadki, gdzie maksimum heatmapy Grad-CAM nie pokrywa się z obiektem lub obiekt jest słabo widoczny.

Potencjalne ulepszenia: 1. **Wykorzystanie progowania heatmapy** - zamiast pojedynczego punktu maksymalnego, użycie kilku punktów powyżej określonego progu pewności (np. 90 percentyla). 2. **Klasteryzacja punktów** - grupowanie punktów z wysoką aktywacją i wybór centroidy jako punktu reprezentatywnego, co zwiększyłoby stabilność lokalizacji. 3. **Inteligentny wybór punktów background** - zamiast stałych rogów obrazu, wybór regionów o najniższej aktywacji w heatmapie Grad-CAM, co zapewniłoby bardziej informacyjne punkty tła. 4. **Wykorzystanie informacji o kształcie** - uwzględnienie informacji o przewidywanej klasie obiektu (np. okrąg vs kwadrat) do dostosowania strategii wyboru punktów. 5. **Ensemble wielu warstw** - wykorzystanie heatmap z różnych warstw sieci neuronowej (np. layer1, layer2, layer3) dla bardziej stabilnych i dokładnych wyników lokalizacji.

5. Podsumowanie

Zadanie zostało wykonane pomyślnie, a wszystkie wymagania zostały spełnione:

- Zaimplementowano metodę Grad-CAM zgodnie z wytycznymi projektu.
- Opracowano dwie strategie automatycznego generowania punktów dla modelu SAM:
 - Strategia 1: Tylko punkty foreground (wybór punktu o maksymalnej aktywacji w heatmapie Grad-CAM)
 - Strategia 2: Punkty foreground i background (dodatkowo punkty w rogach obrazu jako tło)
- Osiągnięto wyniki IoU znacząco przekraczające wymagane 65%:
 - Strategia 1: 72.6% IoU
 - Strategia 2: 72.1% IoU
- Obliczono wszystkie wymagane metryki ewaluacyjne:
 - Hit Rate: 68.4% (strategia 1), 34.3% (strategia 2)
 - Średnia odległość od środka maski: 5.99 px (strategia 1), 23.67 px (strategia 2)
 - Pełne statystyki IoU (średnia, odchylenie standardowe, minimum, maksimum)

Rozwiązanie potwierdza praktyczną użyteczność połączenia metod explainable AI z modelami segmentacji w zadaniach wymagających pełnej automatyzacji. Mimo prostoty zastosowanych strategii wyboru punktów, osiągnięto bardzo dobre wyniki segmentacji, co świadczy o skuteczności podstawowej koncepcji wykorzystania Grad-CAM do automatycznej lokalizacji obiektów.