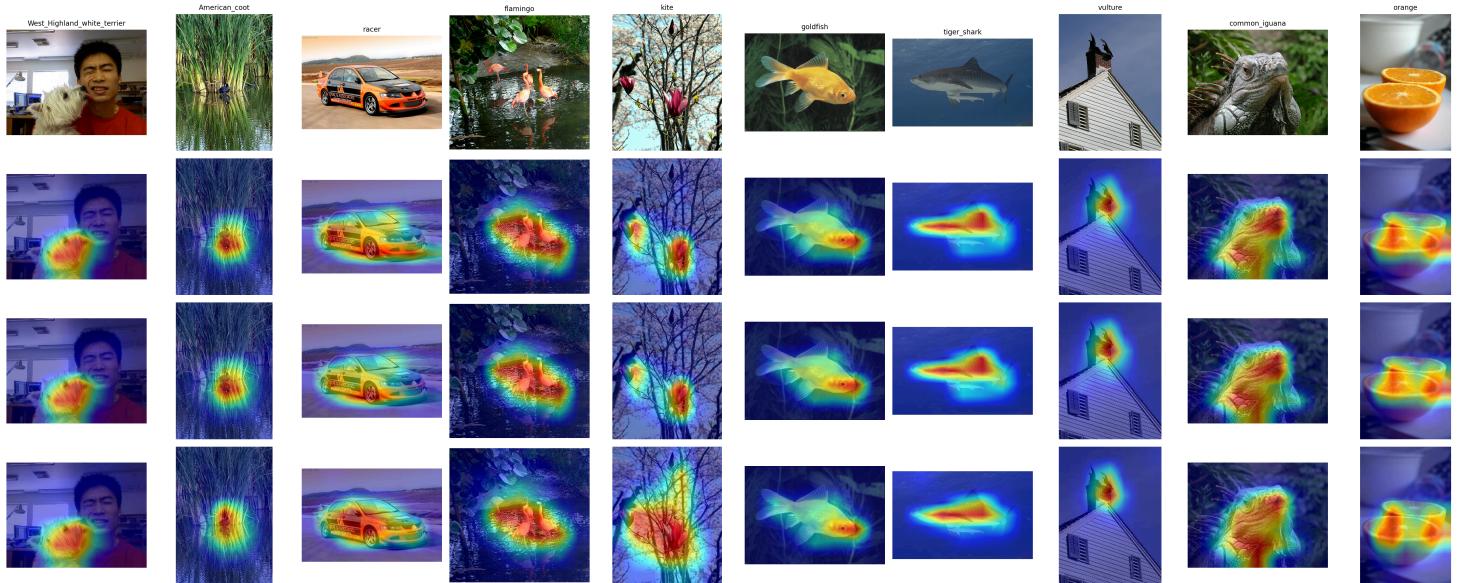


Grad-CAM Analysis Report



Overview

In this task, we applied **Grad-CAM**, **AblationCAM**, and **ScoreCAM** to 10 selected images from the ImageNet dataset using a pre-trained ResNet50 model. The combined 4×10 grid figure shows the original image (top row) and the three CAM variants below.

Note: The second, third, and fourth rows correspond to **Grad-CAM**, **AblationCAM**, and **ScoreCAM**, respectively.

Observations and Analysis

1. Focus on Key Objects

- For most images (e.g., *goldfish*, *tiger_shark*, *orange*), the highlighted areas align well with the main object of interest.
- The model tends to focus on the object's distinctive features (e.g., head of the *West Highland white terrier*, or body of the *flamingo*).

2. Background vs. Foreground

- In simpler images with clean backgrounds (*goldfish*, *orange*), Grad-CAM clearly isolates the object with minimal activation on the background.

- For images with cluttered or complex scenes (e.g., *kite*, *common_iguana*), some background regions also receive attention, indicating potential confusion or reliance on contextual cues.

3. Comparing Grad-CAM, AblationCAM, and ScoreCAM

- The three methods yield **similar activation patterns** across most images, with only subtle variations:
 - **ScoreCAM** tends to produce smoother and more diffused heatmaps.
 - **AblationCAM** highlights slightly smaller and sharper areas.
 - **Grad-CAM** balances both but can be sensitive to gradient noise.
- The differences are most noticeable in images with multiple objects (*kite*, *flamingo*).

4. Robustness and Interpretability

- For objects with clear shapes (*racer*, *goldfish*), all three methods produce consistent heatmaps, showing high confidence in the model's focus.
- For *American_coot*, the activation is tightly centered on the bird's body, demonstrating good localization.

Conclusion

Grad-CAM and its variants effectively highlight the model's attention areas. For the given 10 ImageNet images:

- The highlighted regions mostly correspond to the expected object areas, validating the model's focus.
- Differences between Grad-CAM, AblationCAM, and ScoreCAM are minimal in simple images but slightly pronounced in cluttered scenes.
- Overall, these visualization methods provide valuable interpretability insights into the model's decision-making.