

Task 4 - Comparative Analysis of Grad-CAM and LIME with IoU Scores

This report analyzes and compares two model interpretability techniques, Grad-CAM and LIME, based on their application to ten ImageNet images. The analysis is supported by quantitative Intersection over Union (IoU) scores to precisely measure the degree of agreement between the two methods' explanations.

1. Core Methodological Differences

The fundamental difference between Grad-CAM and LIME dictates the nature of their outputs:

- Grad-CAM (Gradient-weighted Class Activation Mapping)** is a **gradient-based** method. It produces a continuous **heatmap** highlighting the most influential regions for a model's decision by analyzing gradients flowing into the final convolutional layer.
- LIME (Local Interpretable Model-agnostic Explanations)** is a **perturbation-based** method. It identifies important regions by creating and analyzing variations of the input image, resulting in an explanation composed of discrete **image segments** (superpixels).

2. Quantitative Comparison with Intersection over Union (IoU)

We calculated the IoU for the masks generated by both methods for each of the ten images. The results reveal a significant and quantifiable divergence in their explanations.

a. IoU-Based Agreement Analysis

We use the **Intersection over Union (IoU)** metric to measure the spatial agreement between the regions highlighted by both methods. Below is a summary of observations:

Image	IoU	Observations
coot.png	0.1258	Highest agreement – both methods focus on whole coot region.
vulture.png	0.1073	High agreement – both highlight whole vulture region.
iguana.png	0.0832	Moderate agreement on head/crest region.
terrier.png	0.0822	Moderate overlap around snout and ears.
shark.png	0.0728	Agreement on outline and dorsal fin.
racer.png	0.0720	Gradcam focused on chassis, while LIME on hood and tail; moderate overlap.
flamingo.png	0.0673	Low agreement; LIME selects tight areas, Grad-CAM is broad.
orange.png	0.0548	Low overlap due to LIME focusing on top section.
kite.png	0.0354	Very low agreement – spatial divergence in focus areas.
fish.png	0.0014	Extremely low – Grad-CAM focus on head, LIME focused on body.

3. Analysis and Key Insights

- High-IoU examples** like `coot` and `vulture` have clear object boundaries and are centered in the image, likely contributing to the strong spatial alignment.
- Moderate-IoU examples** such as `iguana`, `terrier`, and `shark` show some agreement around distinct features like heads or fins.
- Low-IoU cases**, including `kite` and `fish`, reveal major spatial divergence. These images have multiple small or disconnected regions of interest.

Conclusion

In conclusion, the IoU analysis provides definitive evidence that Grad-CAM and LIME offer distinct and complementary views into model interpretability. Relying on only one method may provide an incomplete picture of the model's decision-making process.