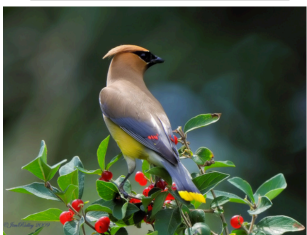


Training Queries



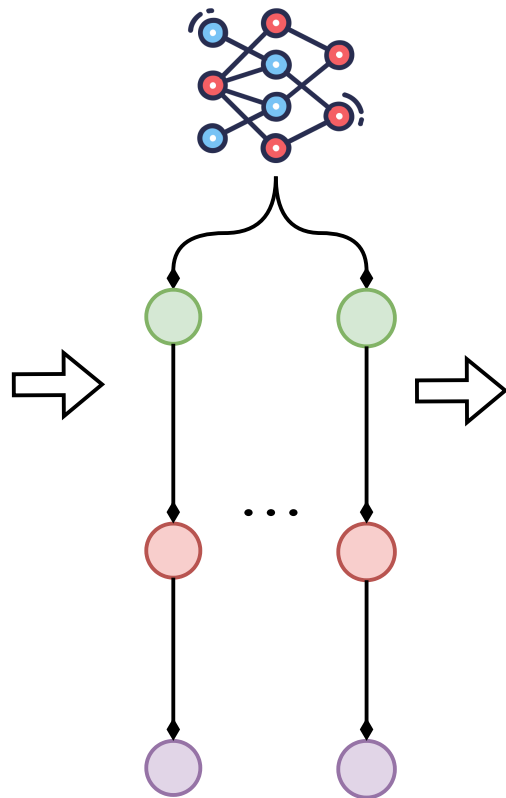
Question: Does the bird in the picture have white underparts and red bill?

Image Caption: The image features a bird perched on a branch with green leaves and red berries.

BBox: $[[152.0, 58.0, 335.0, 328.0]]$

Answer: No.

Multimodal LLM



Capture Global Information



What's in this image?



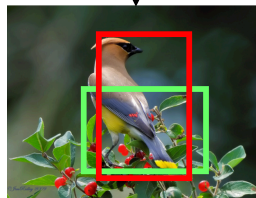
Overall Image Caption



Analysis Critical Regions



Where should I look?

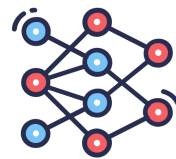


Key Bounding Box



Final Answer

Verifiable Reasoning Patterns



Update Policy



Caption Reward



BBox Reward



Accuracy Reward

Format Reward