

Hallucination

*Lowering identification error between two species of
birds.*

Problem

Llava's Got Issues.

Confirmation bias hallucination in large language models (LLMs) occurs when the model generates responses that align with existing biases or preconceptions, even if they are inaccurate. This happens because LLMs often produce text that fits familiar patterns, reinforcing misinformation and skewed perspectives.



Input: “What color eagle is this?”

Llava: “The color of this eagle is red.”



Input: Llava

What color bird is this?



Black



This bird is brown and white

Our Solution

Our solution focuses on identifying bird species captured in photographs using machine learning techniques. By leveraging the BioCLIP framework, we can accurately analyze and classify bird images, contributing to biodiversity monitoring and conservation efforts. This approach not only enhances species identification but also supports ecological research by providing precise data on bird populations and distributions.



High Level Process



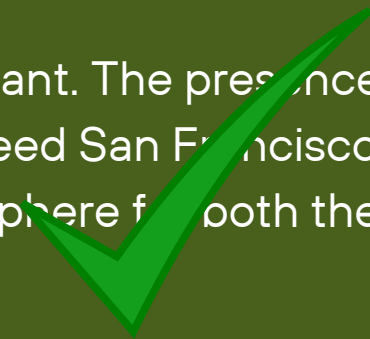
	USE	RESULT
LLAVA	Text GPT for Initial Questioning	This line of questioning will get the initial information that BIOCLIP needs which is to know if the image shows an animal or not.
BIOCLIP	Organism Specific LLM for Species verification	Returns the species of bird in the picture to Llava.

Llava



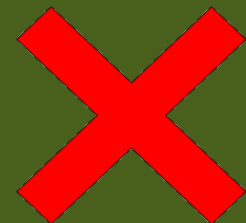
Whats the weather like in San Francisco?

The weather in San Francisco appears to be sunny and pleasant. The presence of a bridge and a city skyline suggests that the location is indeed San Francisco. The sunny weather is likely to contribute to a pleasant atmosphere for both the bird and the people in the city.



What color is the dog?

Yellow

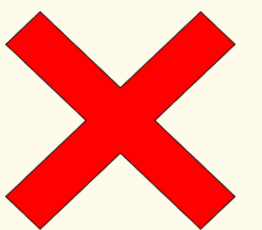


BioClip



Whats the weather like in San Francisco?

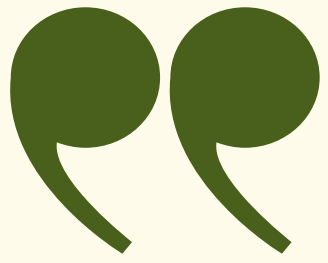
[BioClip does not have any text based reasoning]



What color is the dog?

American Goldfinch

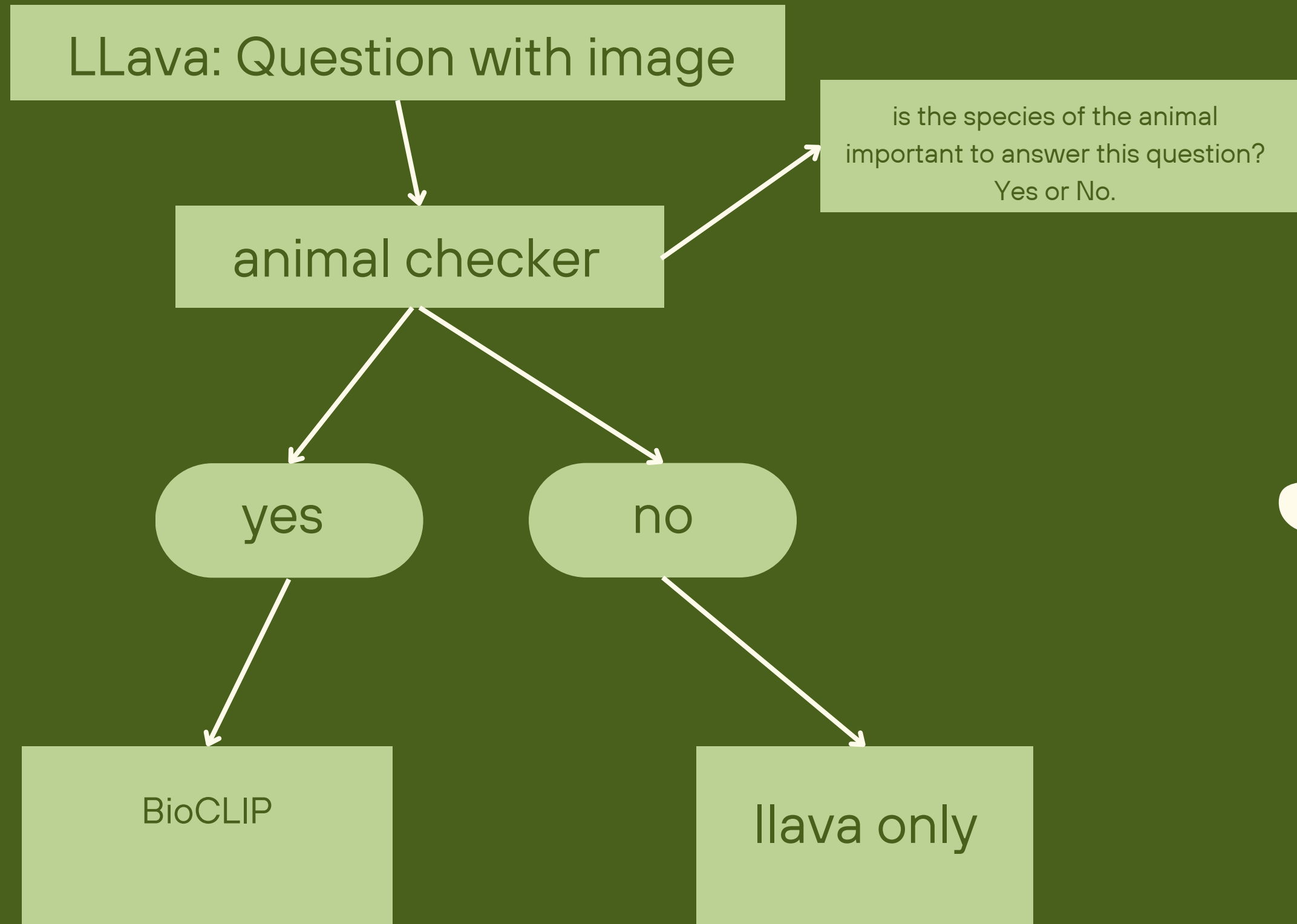




We rigorously benchmark our approach on diverse fine-grained biology classification tasks and find that BIOCLIP consistently and substantially **outperforms existing baselines** (by 16% to 17% absolute). Intrinsic evaluation reveals that BIOCLIP has learned a hierarchical representation conforming to the tree of life, shedding light on its strong generalizability

We comprehensively evaluate BIOCLIP on 10 fine-grained image classification datasets covering animals, plants, and fungi...





Sequence



Demo

Limitations

Data

No geographic data (eg., GPS coordinates) is available to BioClip, because of this it could not provide habitat location of species.

Over-Reliance

One should not rely on the model too much. The predictions cannot replace the work of biologists or scientists but want it assist biologists in their work.

Image Classifier

Many complex tasks can be difficult to specify just through text





- **Image Loading Errors:**

- If the image path is incorrect or the image cannot be opened, the script raises an OSError.

- **Multiple species in one image :**

- The script may fail to accurately handle images with multiple species, leading to incomplete or incorrect identifications.

- **Inconsistent Species Identification:**

- If LLAVA and BioClip provide inconsistent species identifications, the script may need to handle this differently.

FAILURE CASES

- **Merge the models**
 - We want to merge the BioCLIP with LLava so that we can reduce the computational cost and make it more efficient.
- **Improve the feature**
 - Make it more user friendly.
- **Multiple species identifier**
 - Add the object detection to the feature to enable LLava to be useful when dealing with more informative picture.

IMPROVEMENTS

Thank You

References

Radford et al. 2021

<https://huggingface.co/imageomics/bioclip>

https://openaccess.thecvf.com/content/CVPR2024/papers/Stevens_BioCLIP_A_Vision_Foundation_Model_for_the_Tree_of_Life_CVPR_2024_paper.pdf

https://openaccess.thecvf.com/content/CVPR2024/papers/Stevens_BioCLIP_A_Vision_Foundation_Model_for_the_Tree_of_Life_CVPR_2024_paper.pdf

