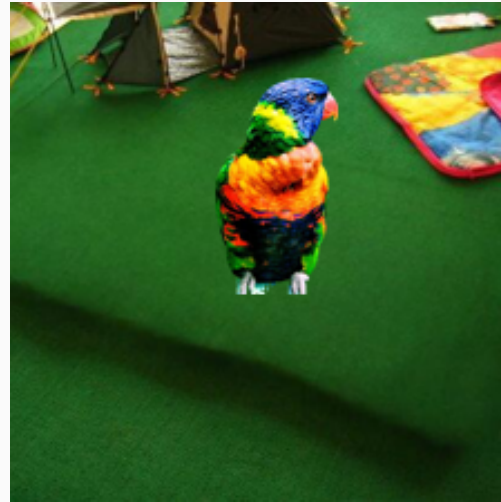


Why does AI call a Parrot a Sleeping Bag?



Current AI:  Sleeping Bag: 69%

Tobias Nauen — YRS 2026

Image classification **supports decisions** in medicine, industry, and biology.

Intro

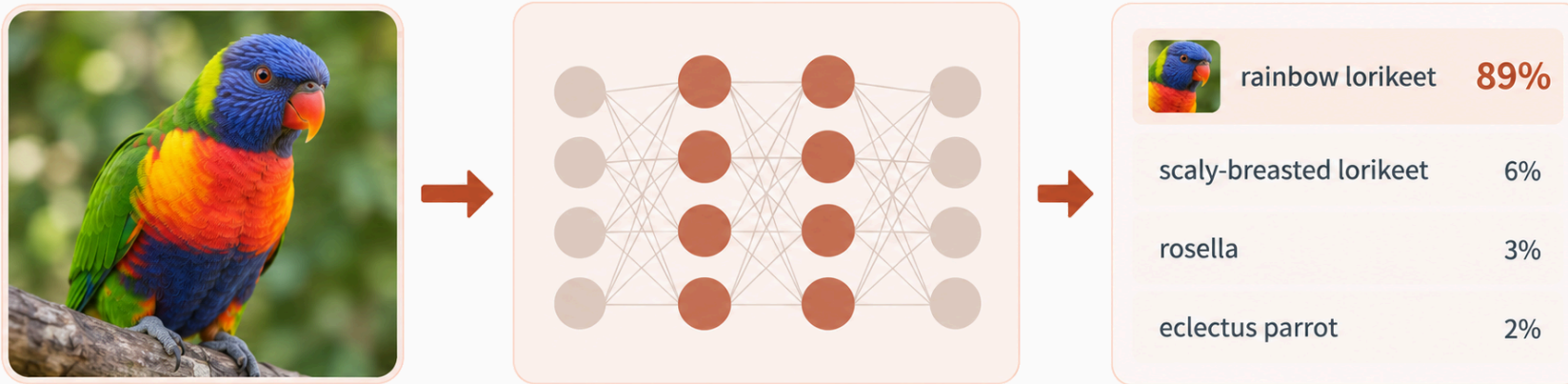
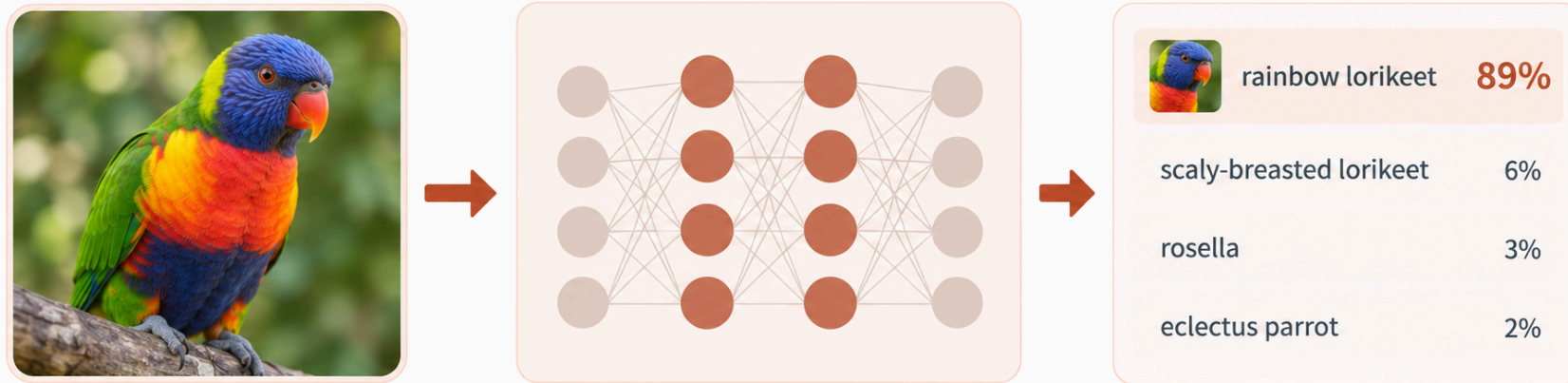


Image classification **supports decisions** in medicine, industry, and biology.

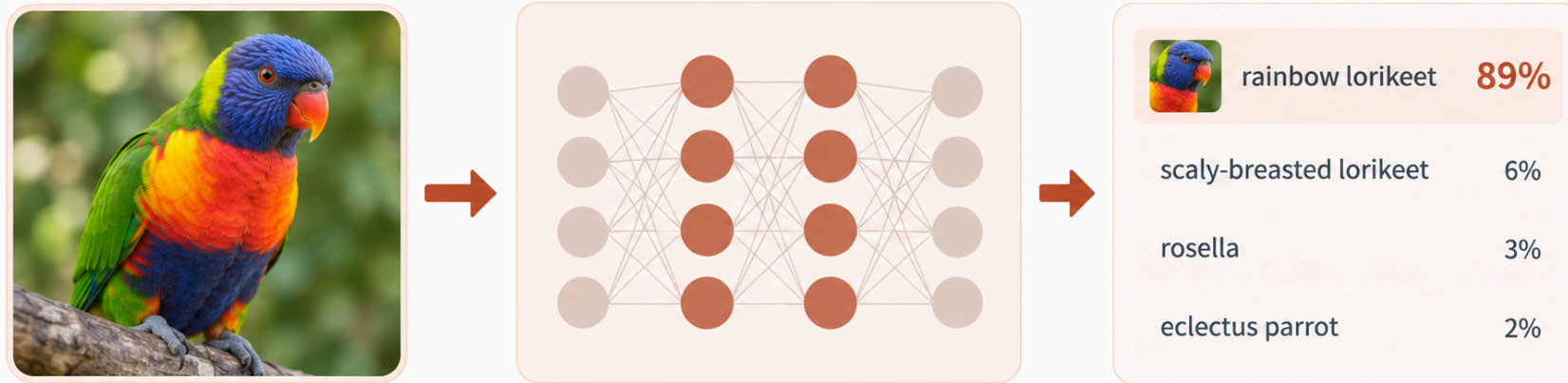
Intro



Medical Diagnosis

Image classification **supports decisions** in medicine, industry, and biology.

Intro



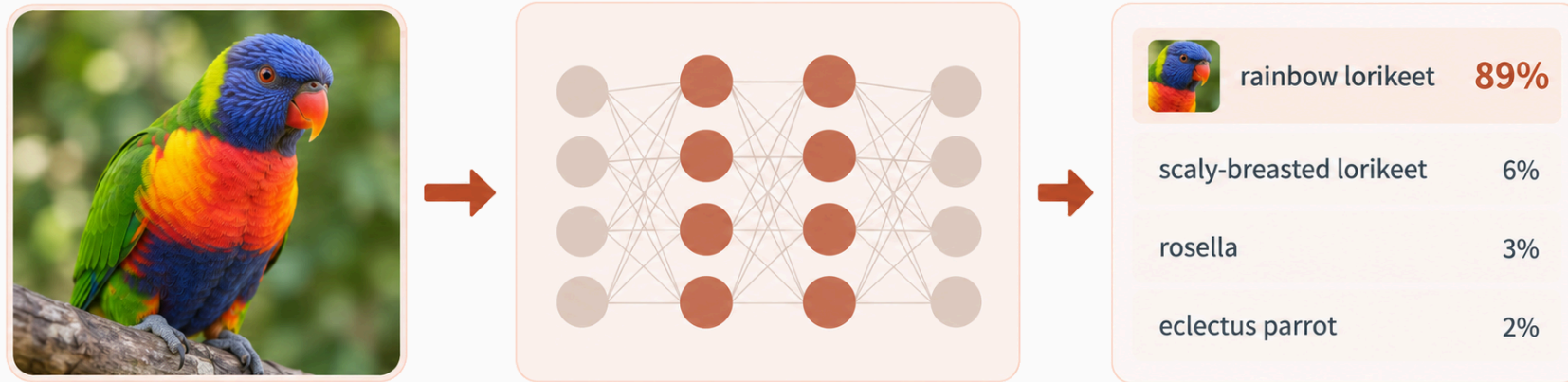
Medical Diagnosis



Quality Control

Image classification **supports decisions** in medicine, industry, and biology.

Intro



Medical Diagnosis



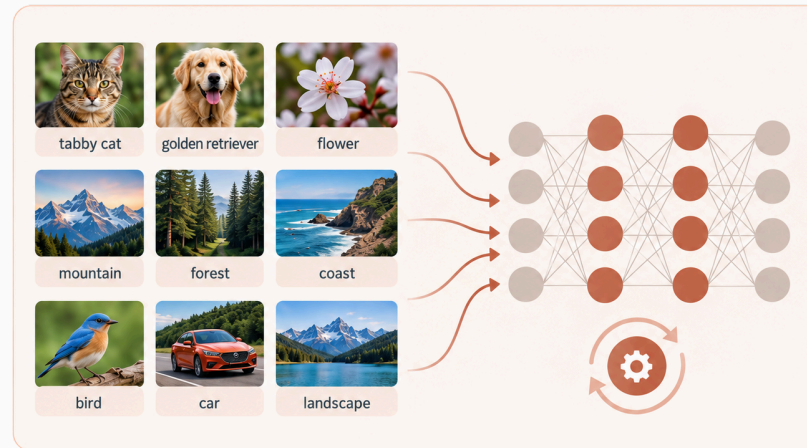
Quality Control



Wildlife Monitoring

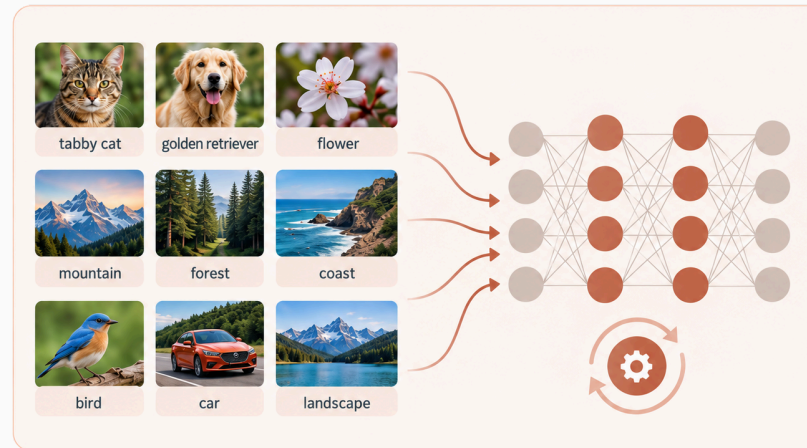
We train models on **data biased** towards what humans take pictures of.

Intro



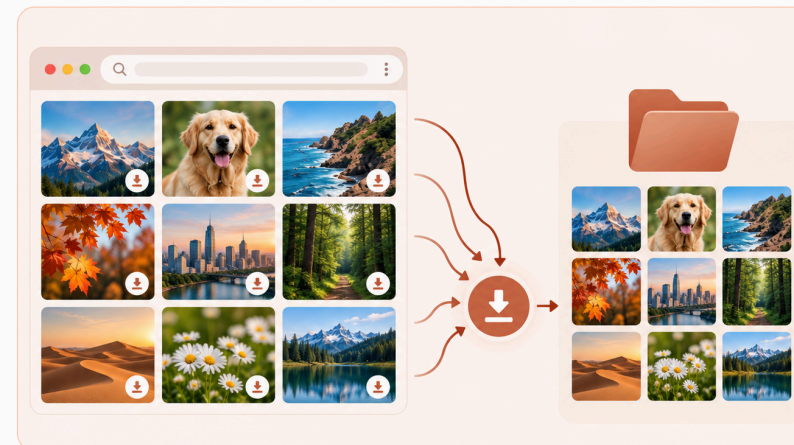
We train models on **data biased** towards what humans take pictures of.

Intro



Images that end up on the internet:

- centered subject
- “normal” context



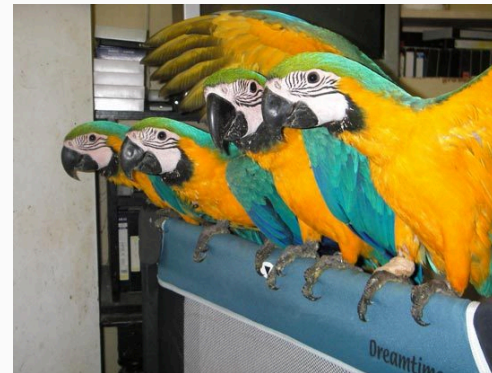
We train models on **data biased** towards what humans take pictures of.

Intro > Examples



We train models on **data biased** towards what humans take pictures of.

Intro > Examples



We train models on **data biased** towards what humans take pictures of.

Intro > Examples



Segmentation models **locate objects** in images for us.

ForAug

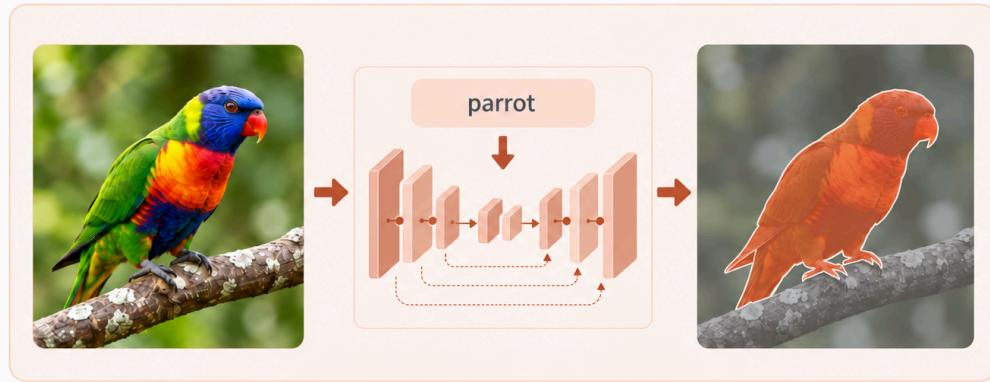


Additional Information



Segmentation models **locate objects** in images for us.

ForAug



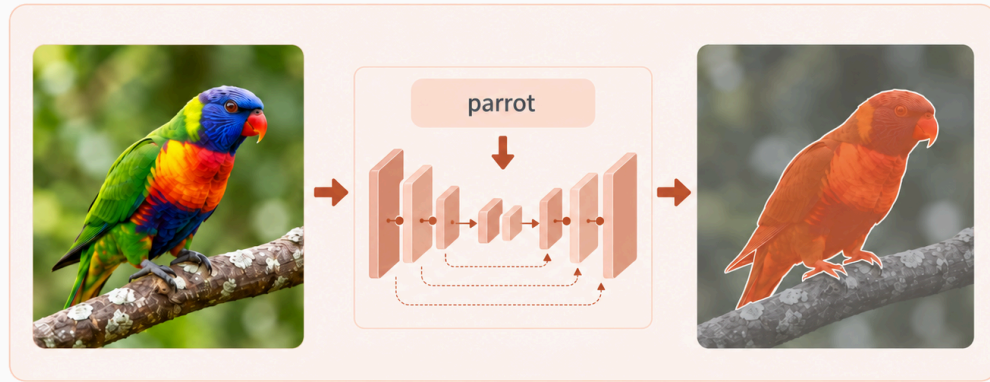
Additional Information

- locate objects in images
- locate broad object type vs. identify specific class



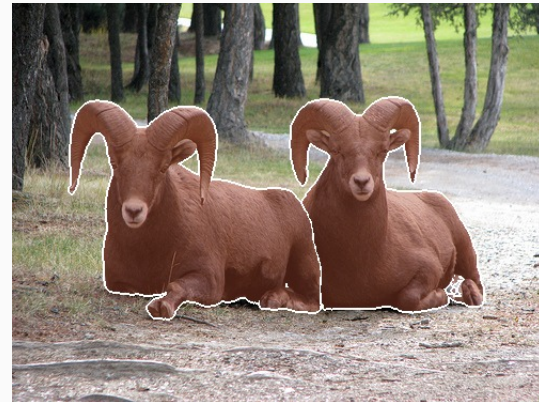
Segmentation models **locate objects** in images for us.

ForAug



Additional Information

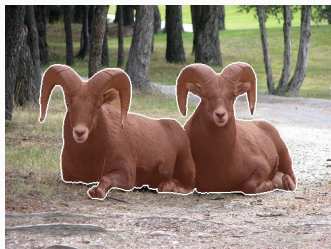
- locate objects in images
- locate broad object type vs. identify specific class



We **extract objects** and fill holes in the backgrounds.

ForAug

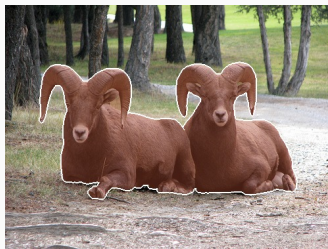
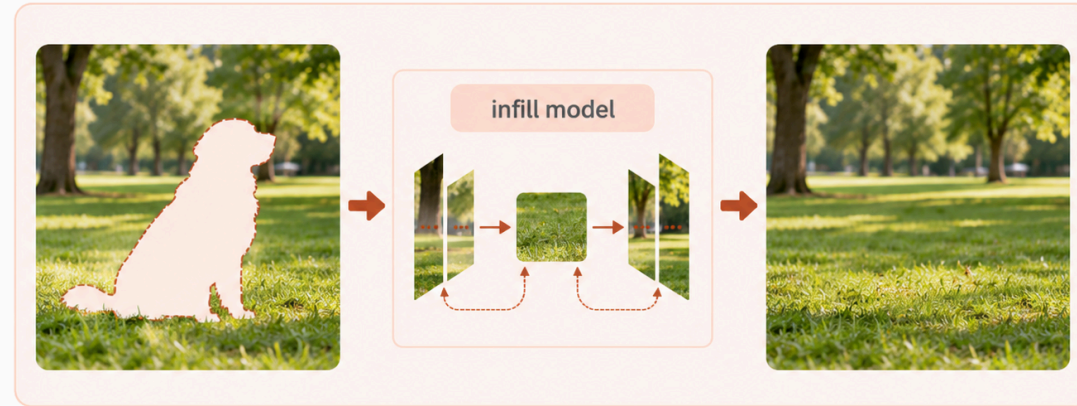
1. Cut out the foreground object



We **extract objects** and fill holes in the backgrounds.

ForAug

1. Cut out the foreground object
2. Fill the hole in the background



We **freely recombine** any object with any background.

ForAug



We **freely recombine** any object with any background.

ForAug



We **freely recombine** any object with any background.

ForAug



We **freely recombine** any object with any background.

ForAug



Vision models perform considerably worse on unexpected backgrounds.

Diagnosis

normal



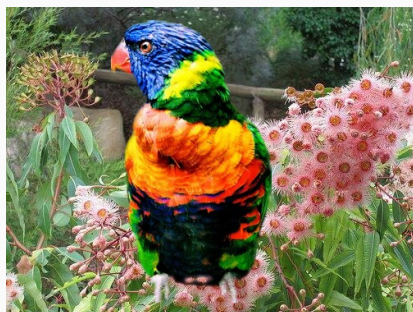
changed



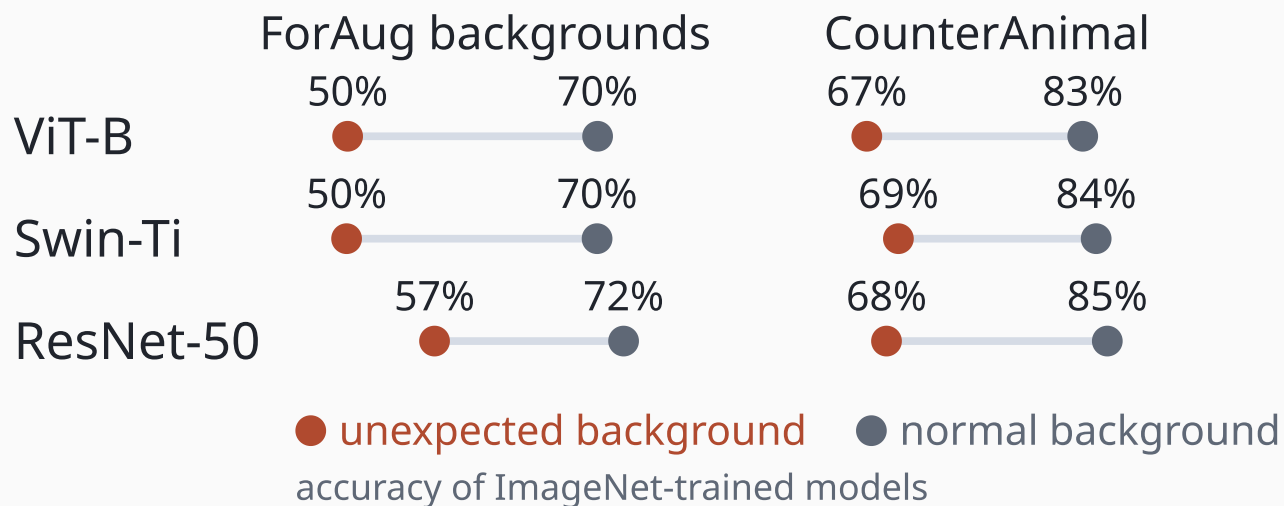
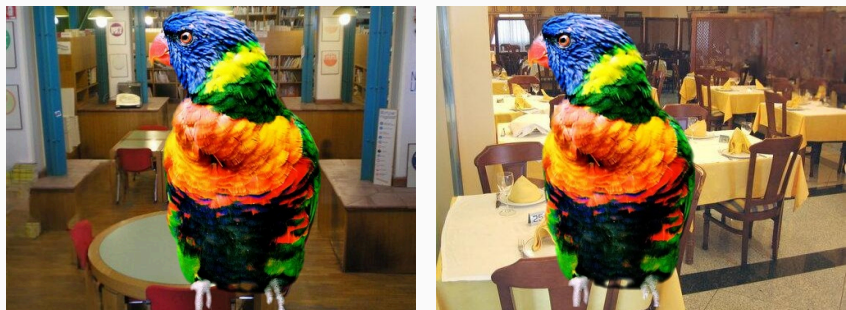
Vision models perform considerably worse on unexpected backgrounds.

Diagnosis

normal



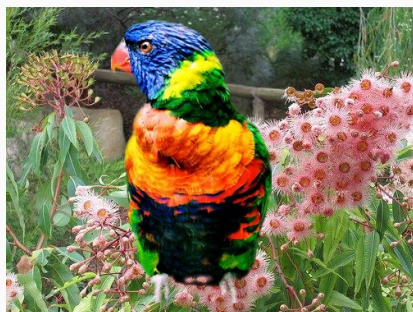
changed



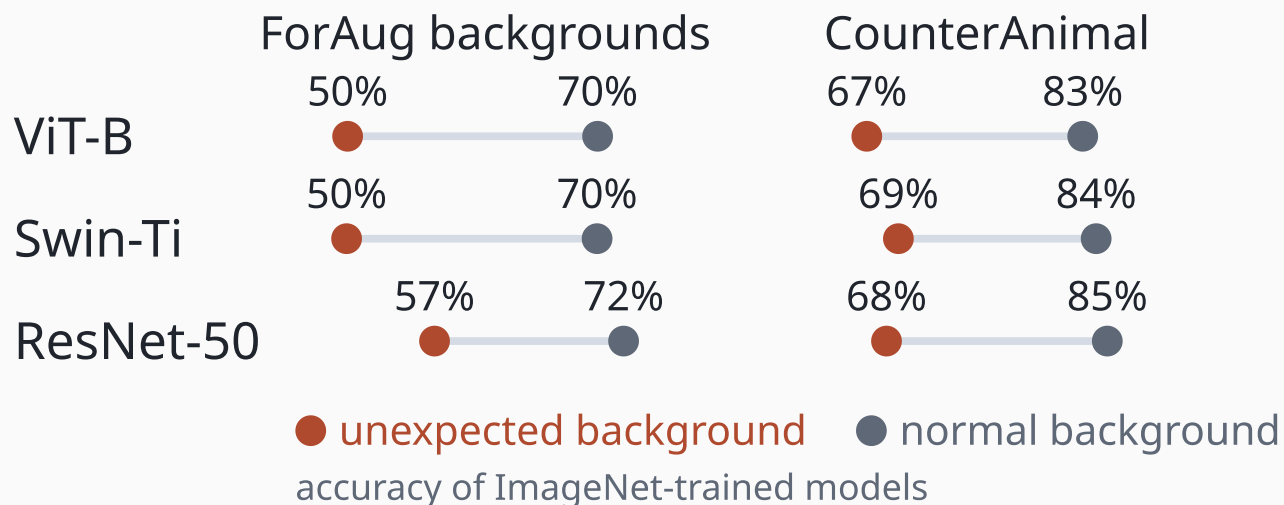
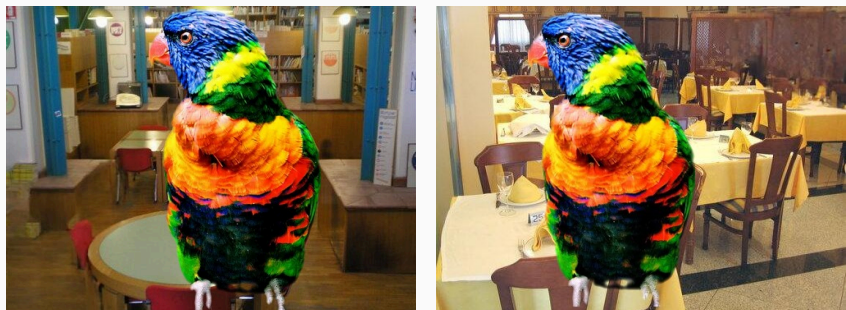
Vision models perform considerably worse on unexpected backgrounds.

Diagnosis

normal



changed



Models lose 15–20% of accuracy when objects are in unexpected contexts.

Vision models perform considerably worse on **non-centered objects**.

Diagnosis

normal



changed



Vision models perform considerably worse on **non-centered objects**.

Diagnosis

normal



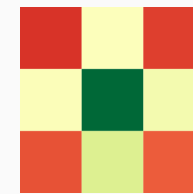
changed



ViT-S



ViT-B



ViT-L



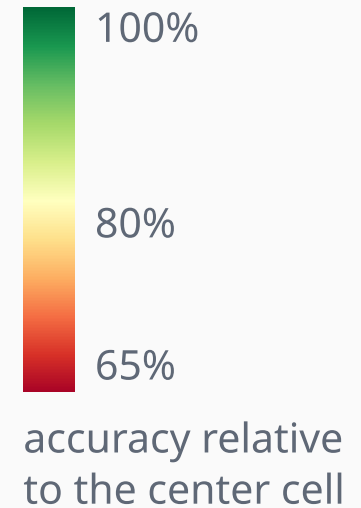
Swin-Ti



Swin-S



ResNet-50



Vision models perform considerably worse on **non-centered objects**.

Diagnosis

normal



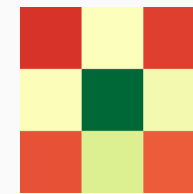
changed



ViT-S



ViT-B



ViT-L



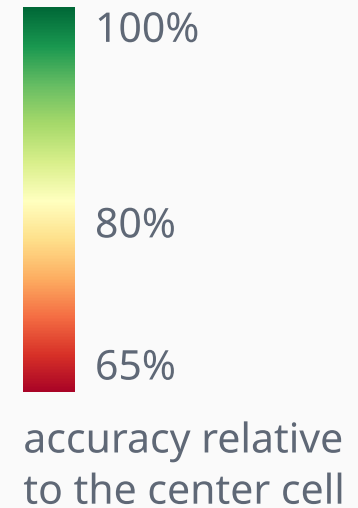
Swin-Ti



Swin-S



ResNet-50



🎯 Objects in corners are hard to detect.

Vision models perform considerably worse on **small objects**.

Diagnosis

normal



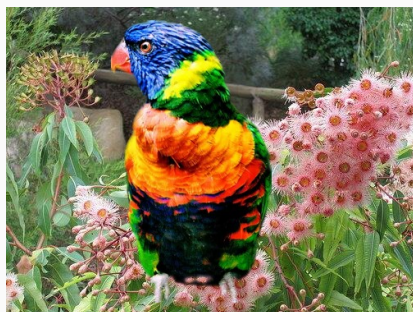
changed



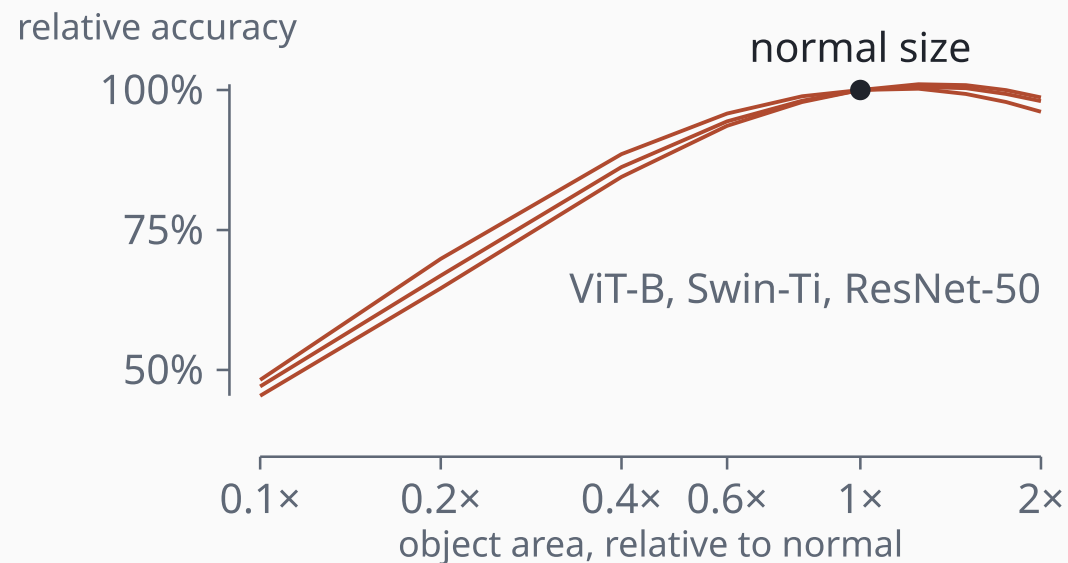
Vision models perform considerably worse on **small objects**.

Diagnosis

normal



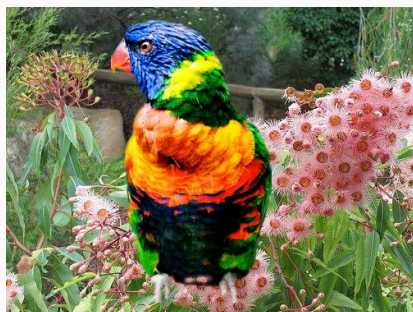
changed



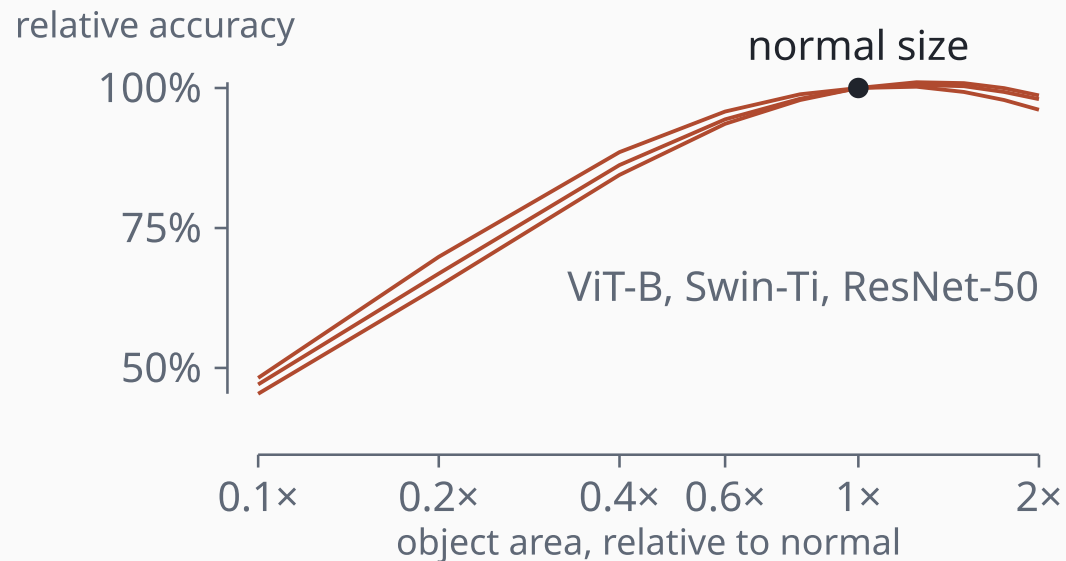
Vision models perform considerably worse on **small objects**.

Diagnosis

normal



changed

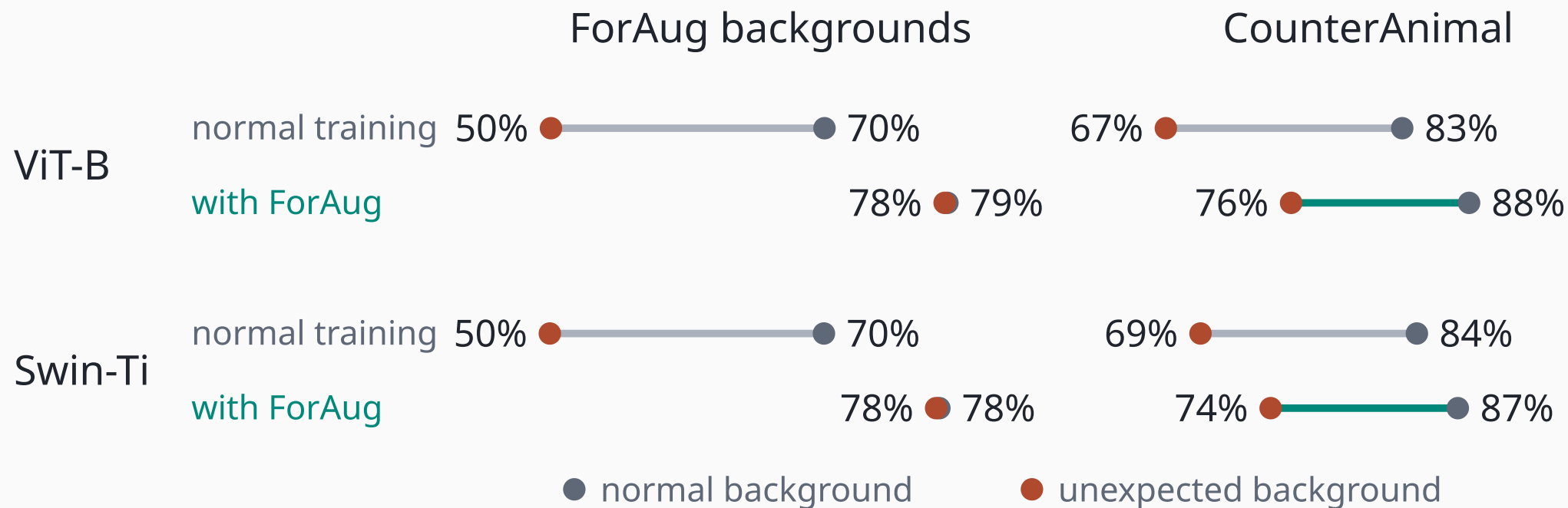


🔍 Accuracy falls quickly on smaller objects.

Training with recombined images **reduces all three biases.**

Solution

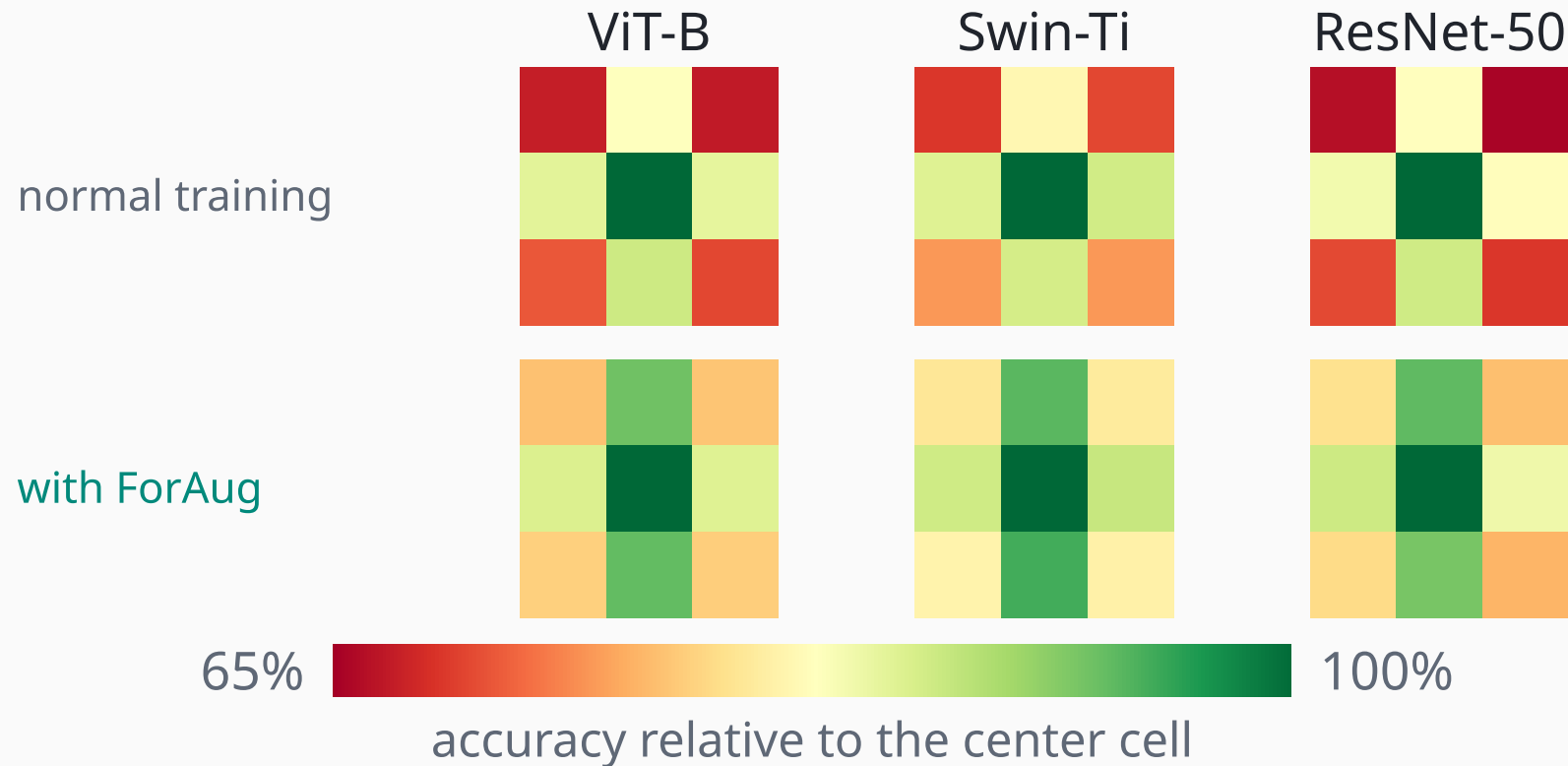
Background bias



Training with recombined images **reduces all three biases.**

Solution

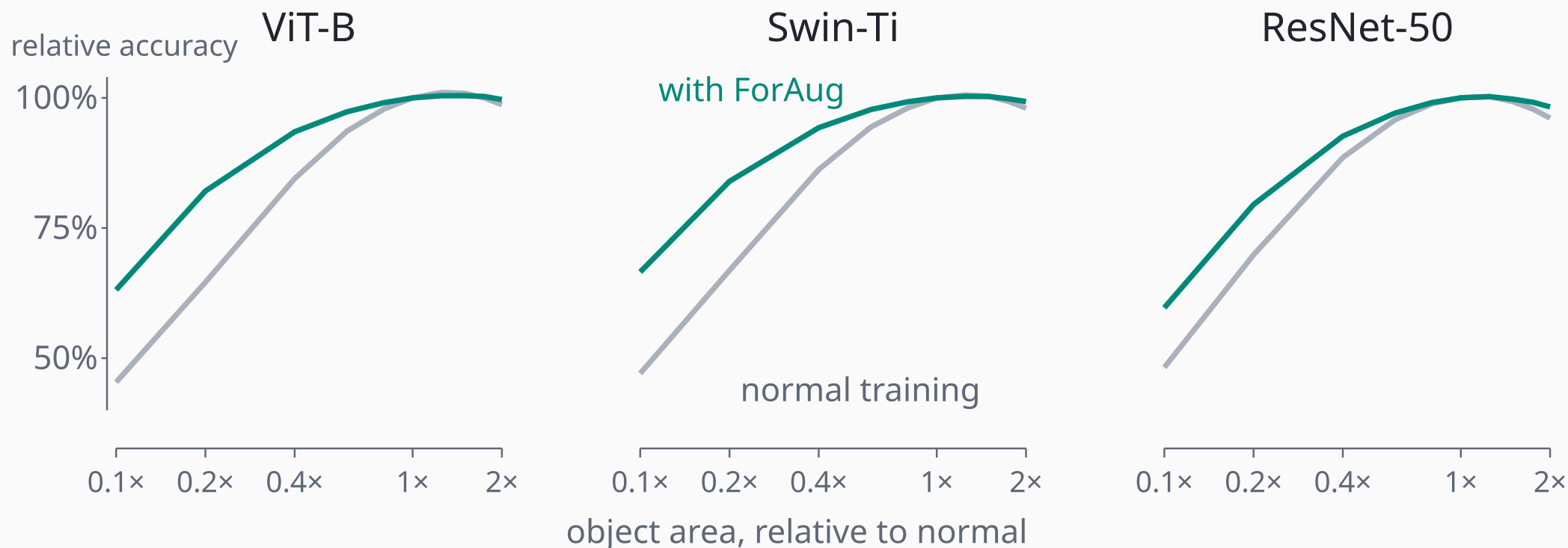
Center bias



Training with recombined images **reduces all three biases.**

Solution

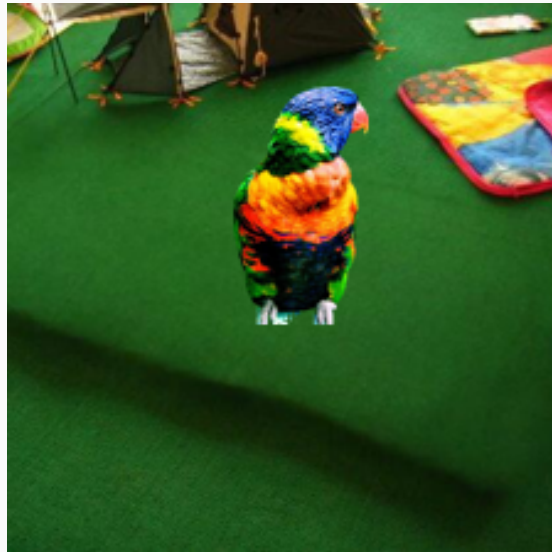
Size bias



Why does AI call a Parrot a Sleeping Bag?

Conclusion

Because our training datasets are biased towards contexts where images are available.

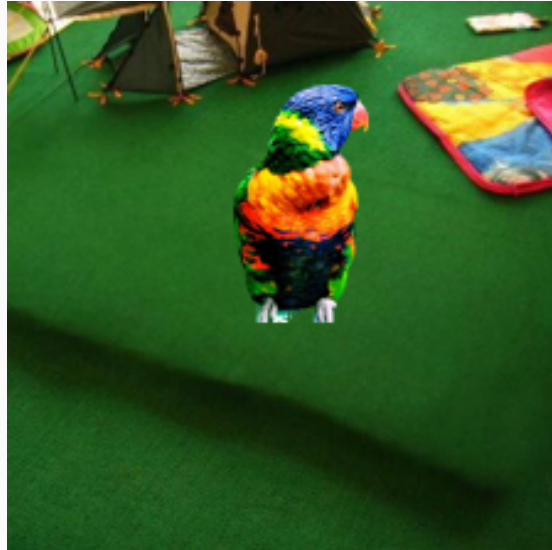


Current AI:  Sleeping Bag: 69%

Why does AI call a Parrot a Sleeping Bag?

Conclusion

Because our training datasets are biased towards contexts where images are available.



We can fix some of these biases by utilizing additional information.

Current AI:  Sleeping Bag: 69%

With ForAug:  Lorikeet: 81%

Naming the exact class requires fine-grained expertise.



What is the correct class for this image?

A: Greater Swiss Mountain Dog

B: Tibetan Mastiff

C: Bernese Mountain Dog

D: Leonberger

Finding the object is easy.



Where in the image is a Bernese Mountain Dog, Dog?

Finding the object is easy.



Where in the image is a Bernese Mountain Dog, Dog?

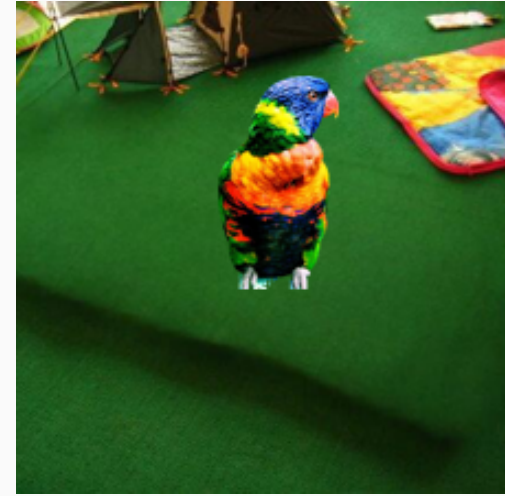
The lorikeet on a sleeping-bag background.



Original



Background



Composite

Predictions on the composite

ViT-B on ImageNet: Sleeping Bag, 69%

ViT-B with ForAug: Lorikeet, 81%

The lorikeet on a **sleeping-bag** background.



Original



Background



Composite

Predictions on the composite

ViT-L on ImageNet: **Sleeping Bag, 74%**

ViT-L with ForAug: **Lorikeet, 84%**

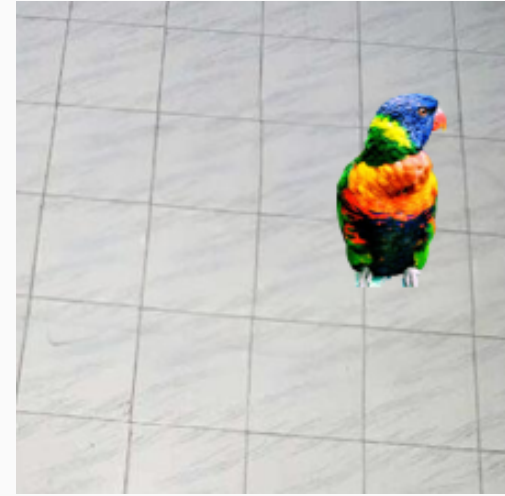
The lorikeet on a **sleeping-bag** background.



Original



Background



Composite

Predictions on the composite

ViT-B on ImageNet: **Sleeping Bag, 78%**

ViT-B with ForAug: Lorikeet, 81%