CV / VLM

Unit 2: Introduction to Object Detection (OD)

TIL-AI
TODAY I LEARNED AI

2.4.2

Handling Challenges

Data augmentation strategies for object detection

TIL-AI

Data Augmentation

- Given the high cost associated with obtaining + annotating training data for object detection, we can explore other ways of creating training data to improve performance

- Data augmentation means to artificially create more training data to improve the model's ability to generalize on difficult conditions to such scenarios
  - e.g. occluded objects and small objects

- Incorporating data augmentation techniques can improve a model's robustness in the real world

Example Python library for image augmentation: albumentations

Data augmentation strategies for OD
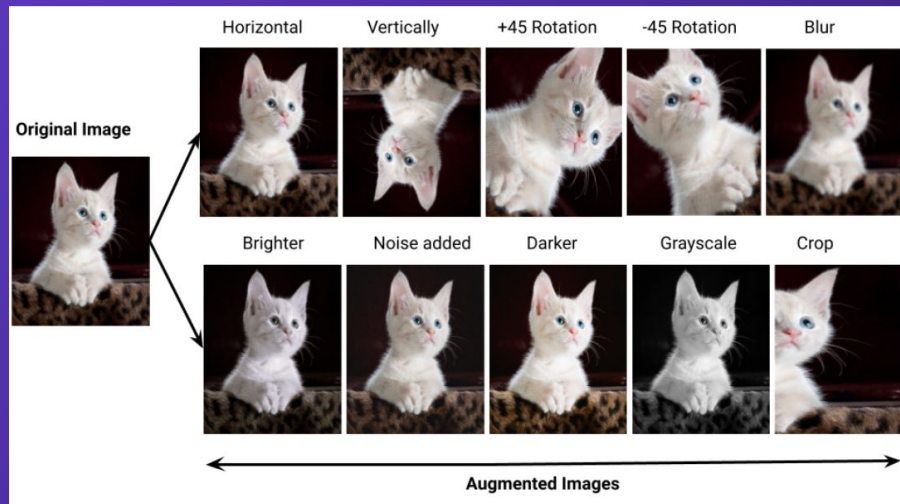## Common operations

Geometric Transformations:

- Flipping (Mirroring)
- Rotation
- Scaling

Color Space Augmentations:

- Color Jittering
- Color Randomization

Other Techniques:

- Noise Injection
- Blurring

Create Augmented Images | Roboflow Docs
Commercial applications

Data augmentations for Computer Vision - TagX - Medium

Data augmentation strategies for OD
Occlusion and Small Objects

To enhance model's ability to discern occluded or small objects, the following techniques can help with enhancing training data.

- **CutMix:** A section of image is randomly cut out and replaced with a section from a different image.
- **Random Erasing:** A random section of image is erased and replaced with noise
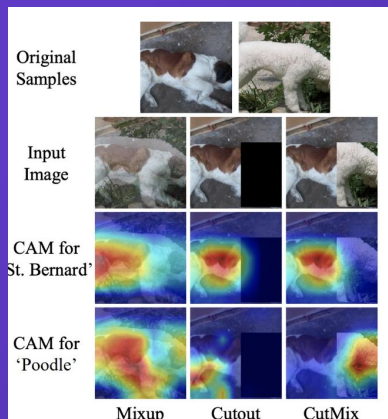

CutMix
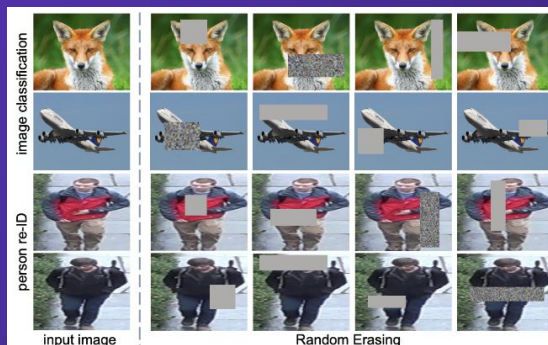
Example of Random Erasing Data Augmentation Zhun Zhong 2017

Factors for Augmentations

Domain expertise

- For instance, for satellite images, cropping, rotation and scaling would be sensible, as it would not introduce distortion

- However, for medical imaging applications, color transformation, grid distortions and elastic transformations would be helpful
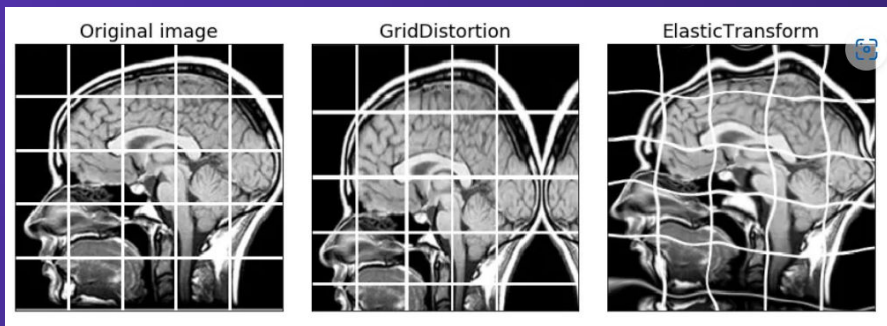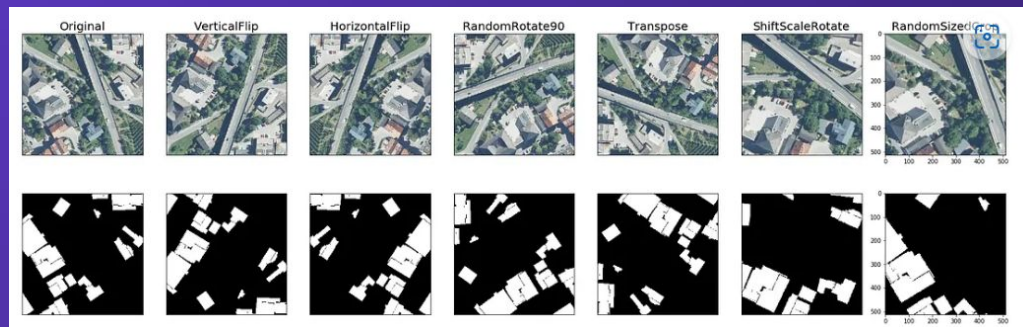
Factors for Augmentations (cont.)

Common Sense

- There is such a thing as too much image augmentation.

- Too much augmentation might lead the model to learn the augmented pattern instead of the original data, thereby harming generalization.

- If the augmented images don't make sense to humans, we wouldn't expect the model to perform well either.

This would be too excessive.