



# HUST

**ĐẠI HỌC BÁCH KHOA HÀ NỘI**  
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

ONE LOVE. ONE FUTURE.





ĐẠI HỌC  
BÁCH KHOA HÀ NỘI  
HANOI UNIVERSITY  
OF SCIENCE AND TECHNOLOGY

# Các kỹ thuật xử lý dữ liệu trong thị giác máy tính

ONE LOVE. ONE FUTURE.

- Image Processing
  - Geometric Transformation
  - Color Transformation
  - Noise Reduction
- Data Augmentation
- Ứng dụng trong bài toán thực tế

- Các phép biến đổi hình học là tập hợp các phép biến đổi hình ảnh từ một hình dạng này sang một hình dạng khác thông qua việc làm thay đổi phương, chiều, góc, cạnh mà không làm thay đổi nội dung chính của bức ảnh

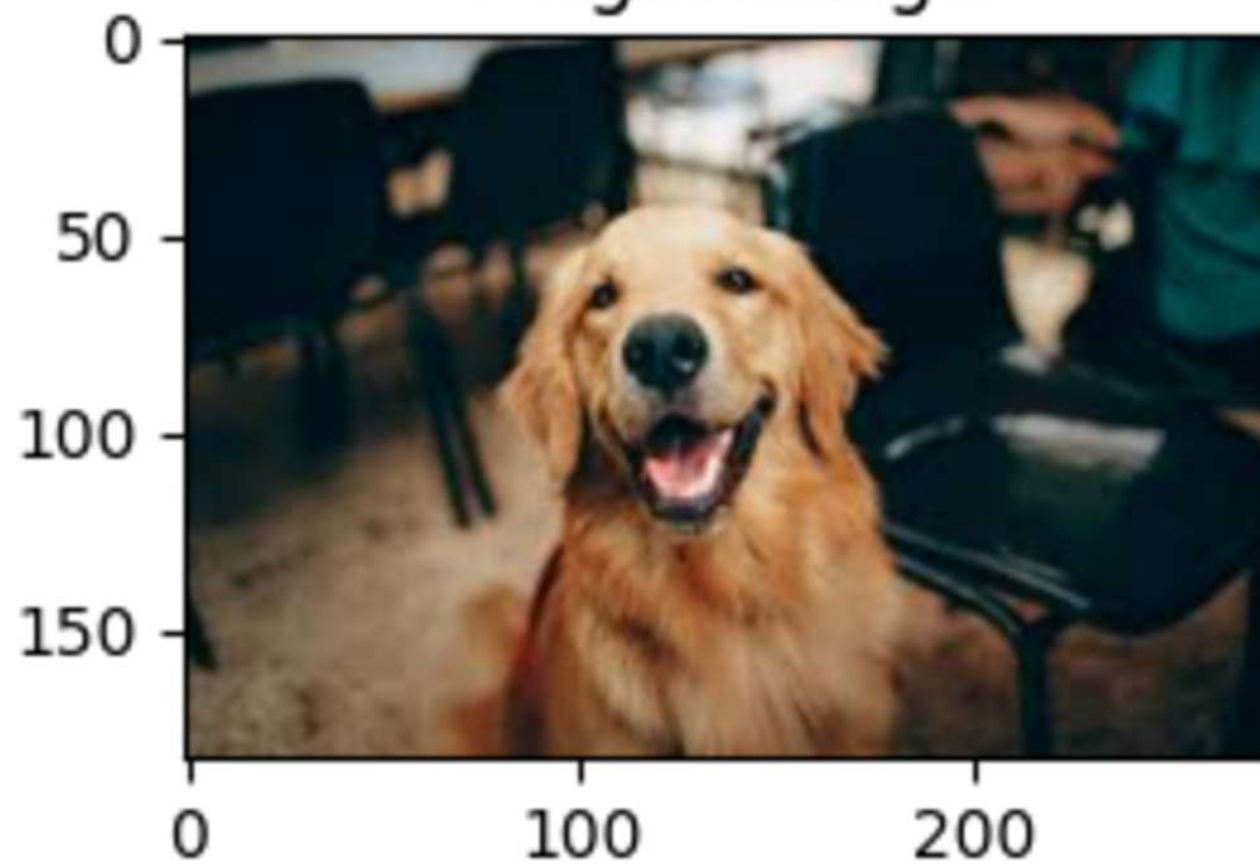
$$T(x, y) = M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{bmatrix}$$



# Scale

$$T(x, y) = M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_1 x \\ a_2 y \end{bmatrix}$$

Origin Image

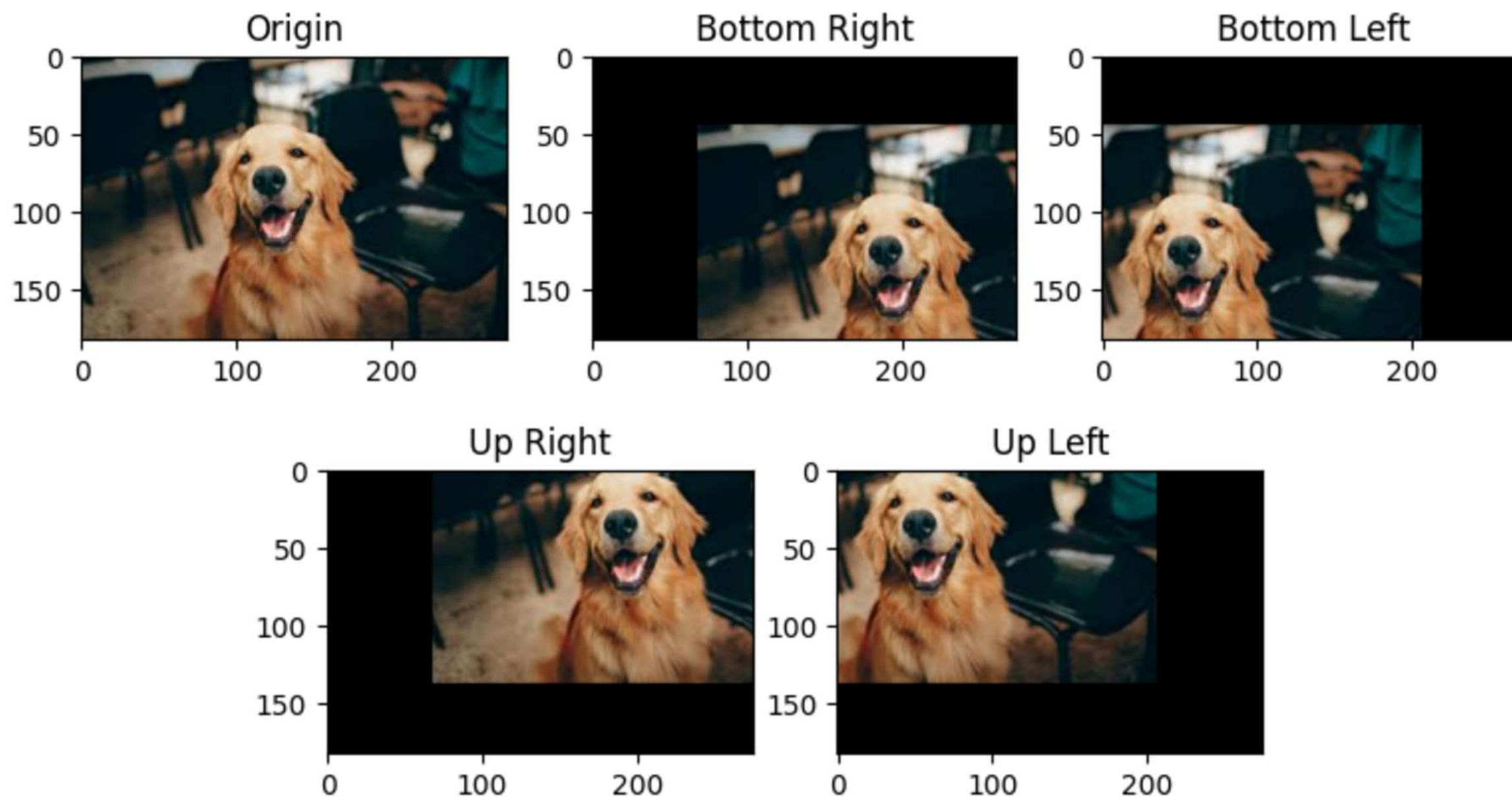


Scale Image



# Translation

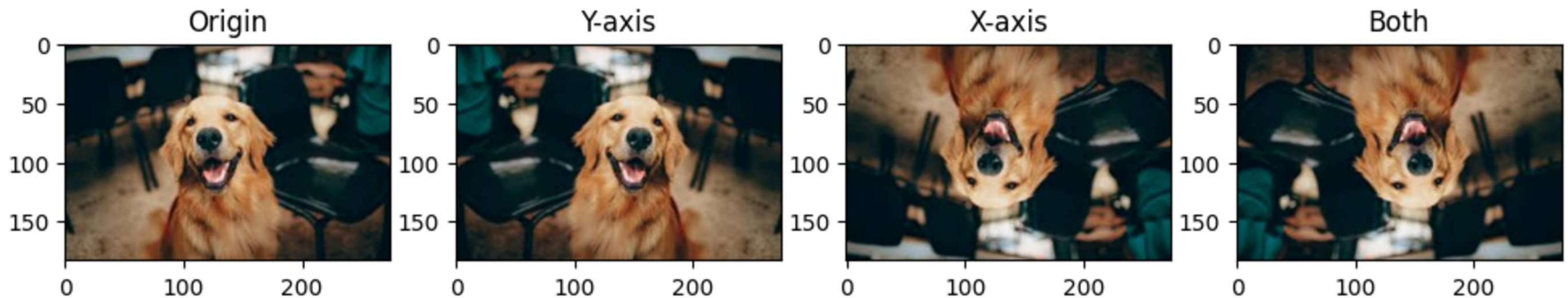
$$T(x, y) = M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + t_x \\ y + t_y \\ 1 \end{bmatrix}$$





# Reflection

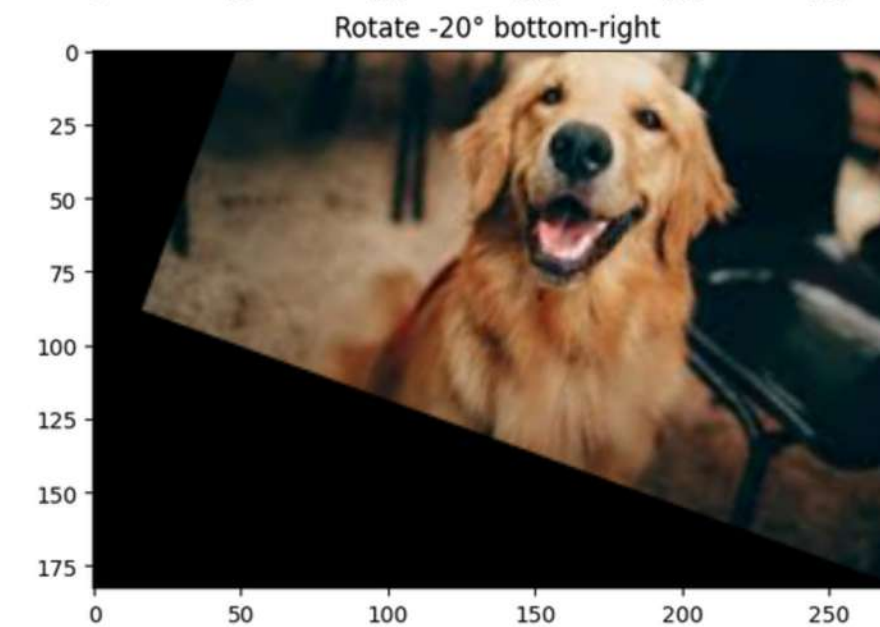
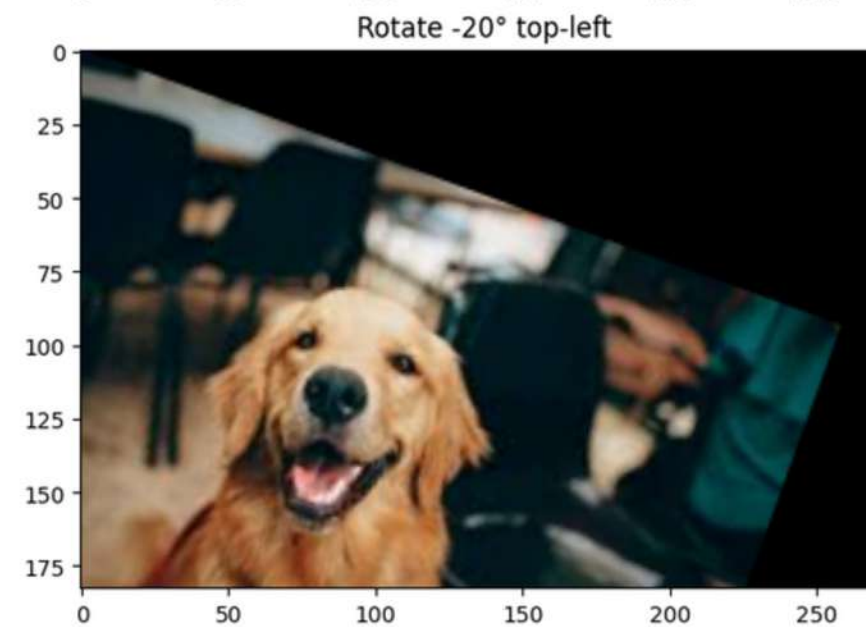
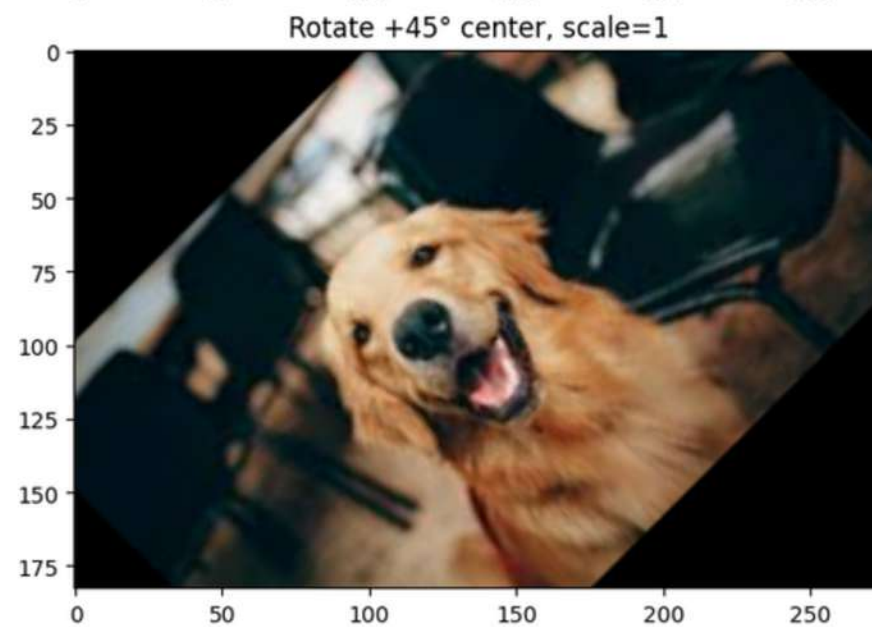
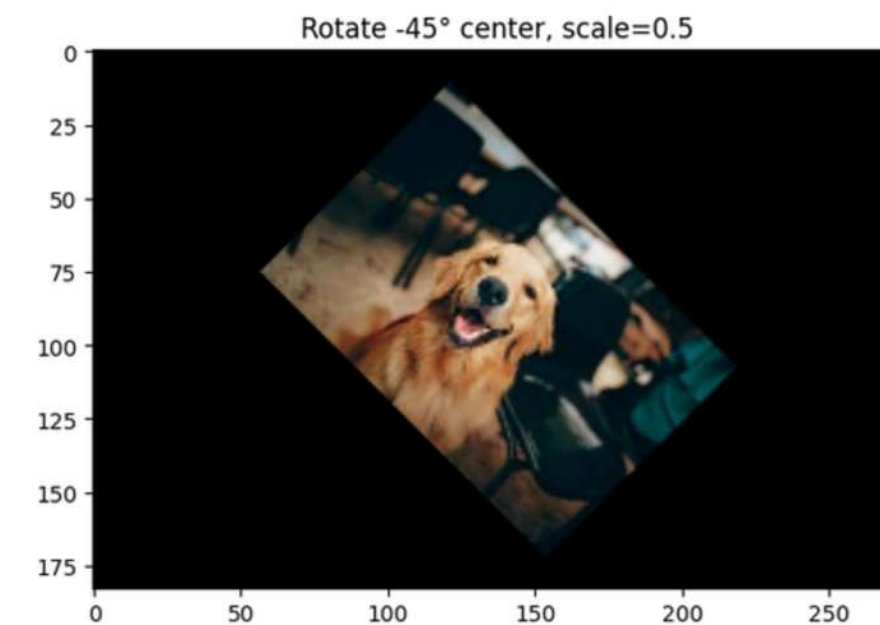
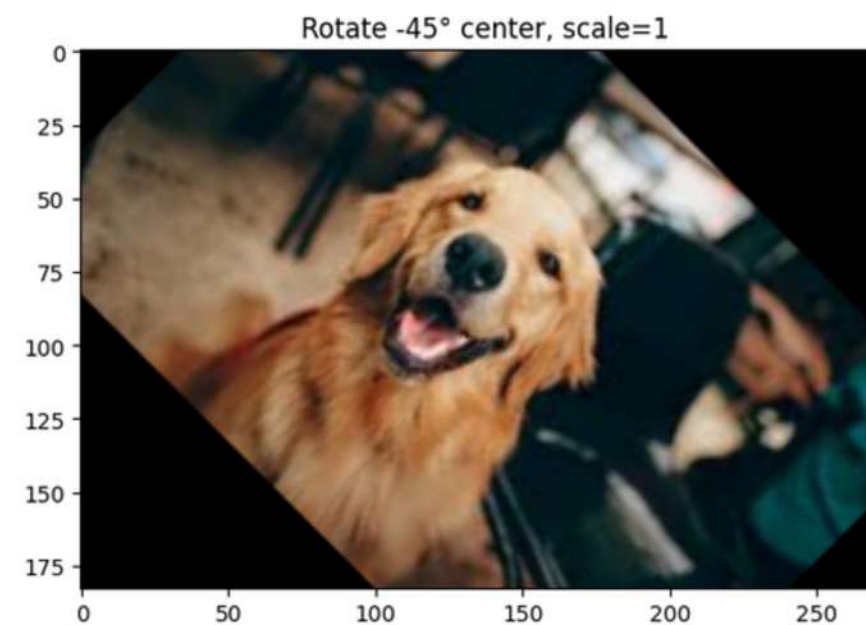
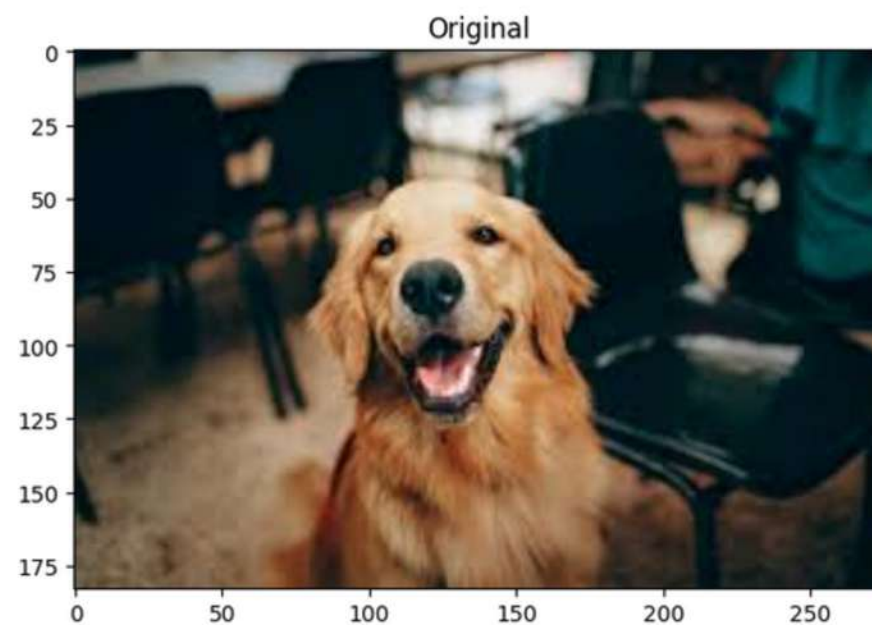
$$T(x, y) = M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 & X \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} X - x \\ y \\ 1 \end{bmatrix}$$





# Rotation

$$T(x, y) = M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \cos(\theta) - y \sin(\theta) \\ x \sin(\theta) + y \cos(\theta) \end{bmatrix}$$



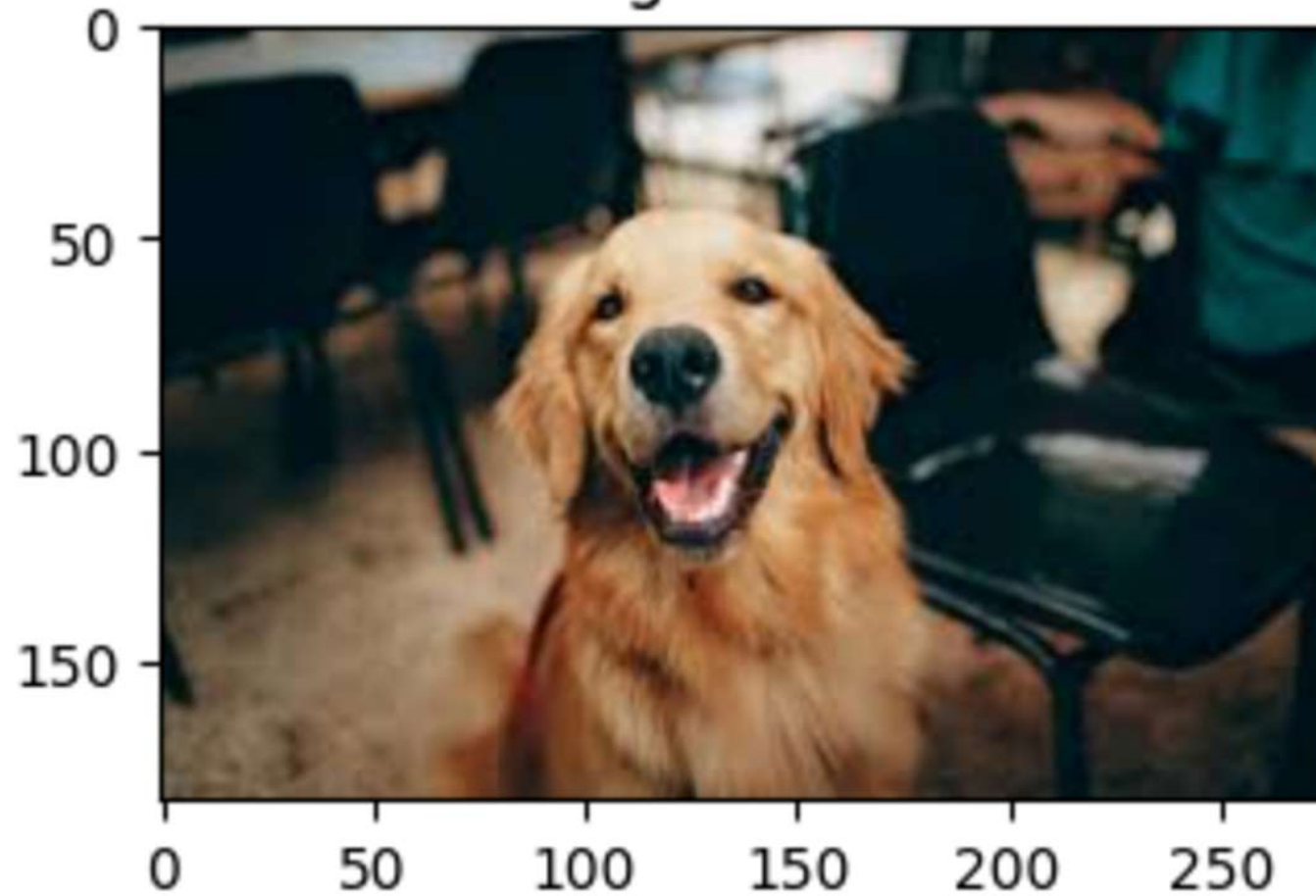


- Các phép biến đổi màu sắc là các phép toán nhằm thay đổi hoặc biểu diễn lại thông tin màu sắc của ảnh từ một không gian màu này sang một không gian màu khác, hoặc biến đổi các giá trị màu để phục vụ một mục đích cụ thể.
- Các không gian màu thường gặp là RGB, Gray, HSV và Lab

# RGB to Gray

$$T(R, G, B) = M \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = [0.299R + 0.587G + 0.114B]$$

Original RGB



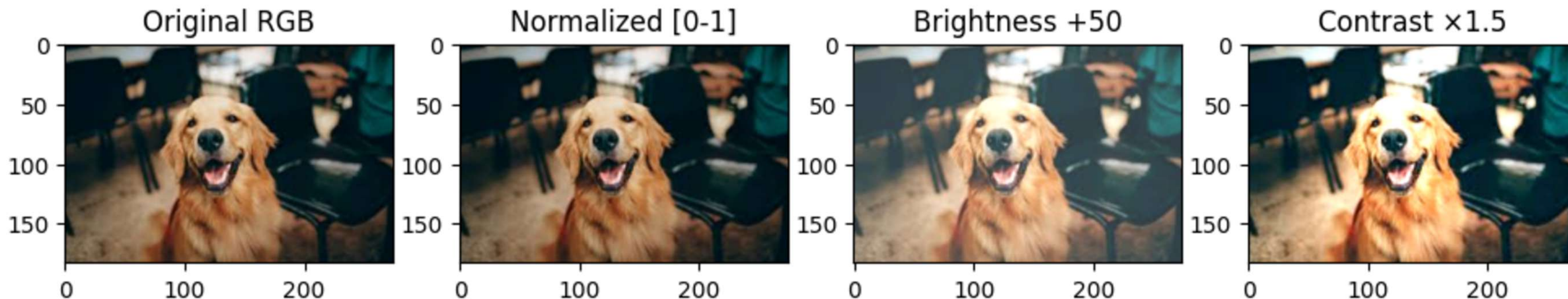
Gray





# Intensity Scaling

$$T(R, G, B) = \begin{bmatrix} a & 0 & 0 & b \\ 0 & a & 0 & b \\ 0 & 0 & a & b \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix} = \begin{bmatrix} aR + b \\ aG + b \\ aB + b \\ 1 \end{bmatrix}$$

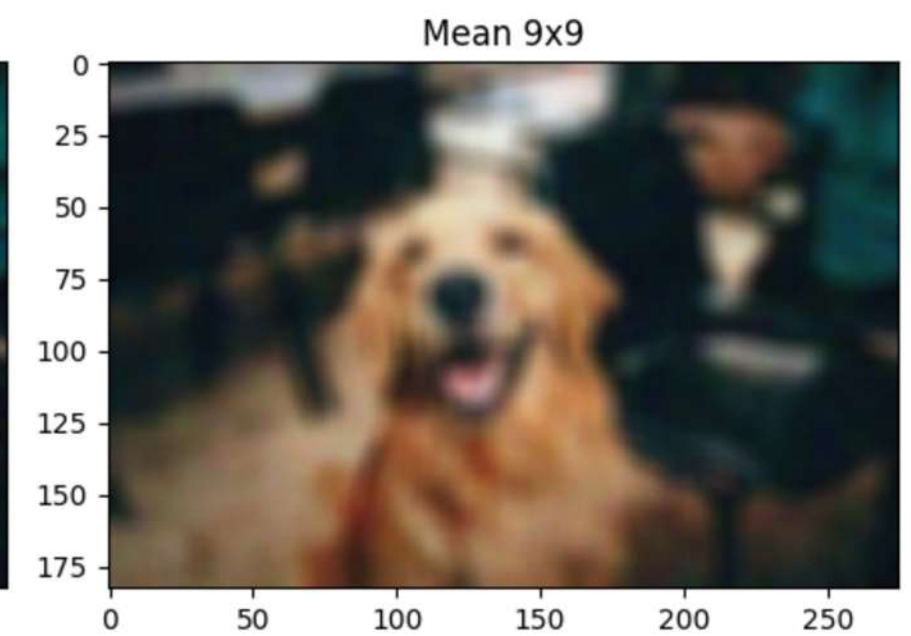
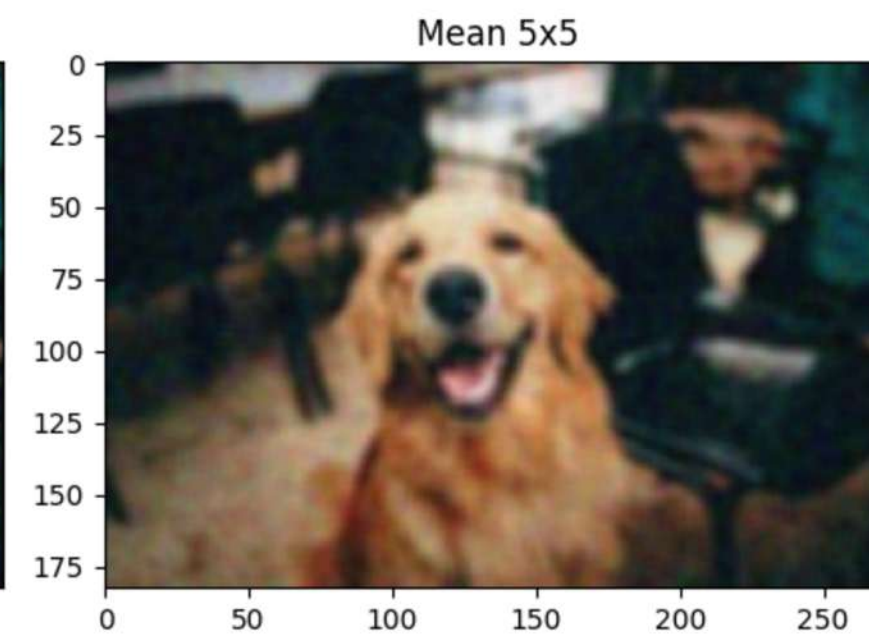
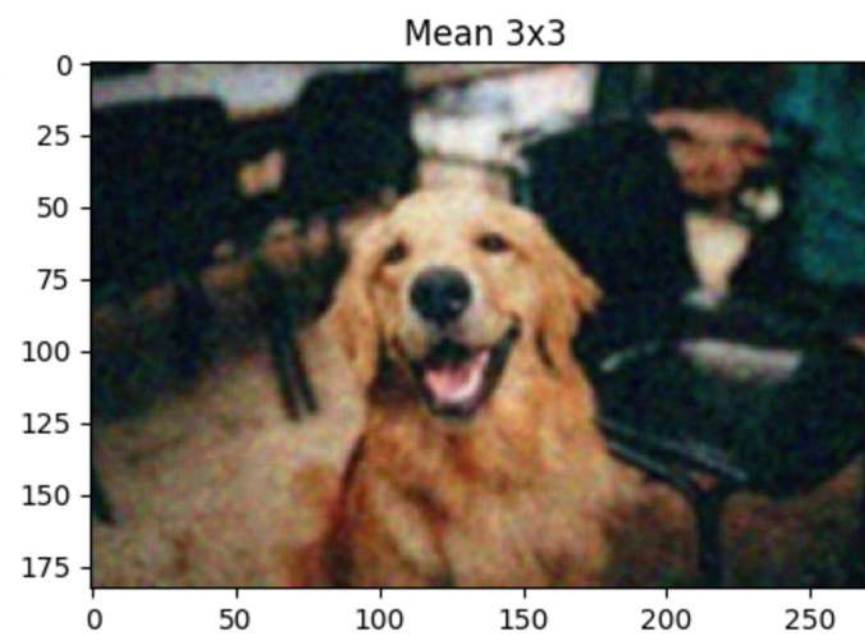
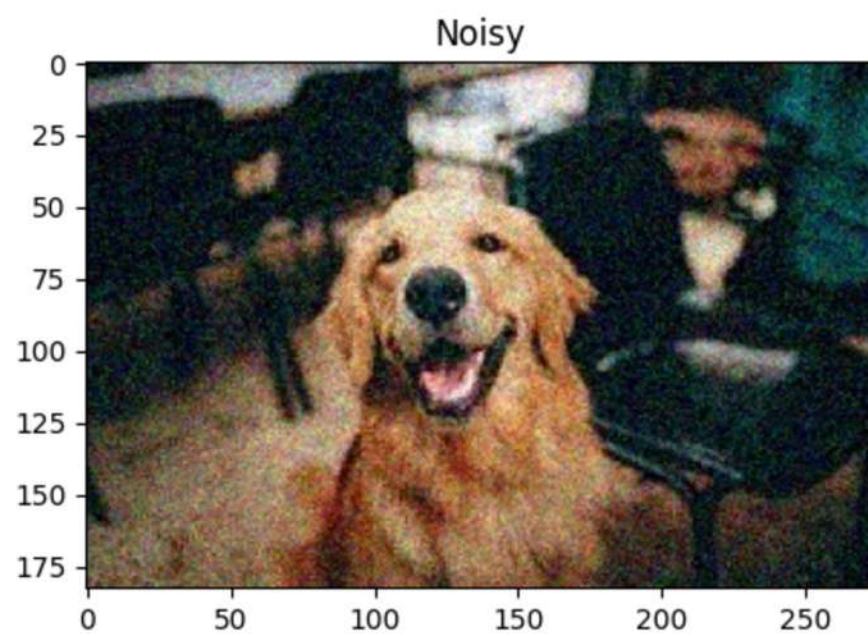


- Noise Reduction là tập hợp các kỹ thuật xử lý ảnh nhằm giảm hoặc loại bỏ nhiễu trong ảnh, trong khi vẫn giữ lại càng nhiều chi tiết quan trọng càng tốt.



# Mean Filter

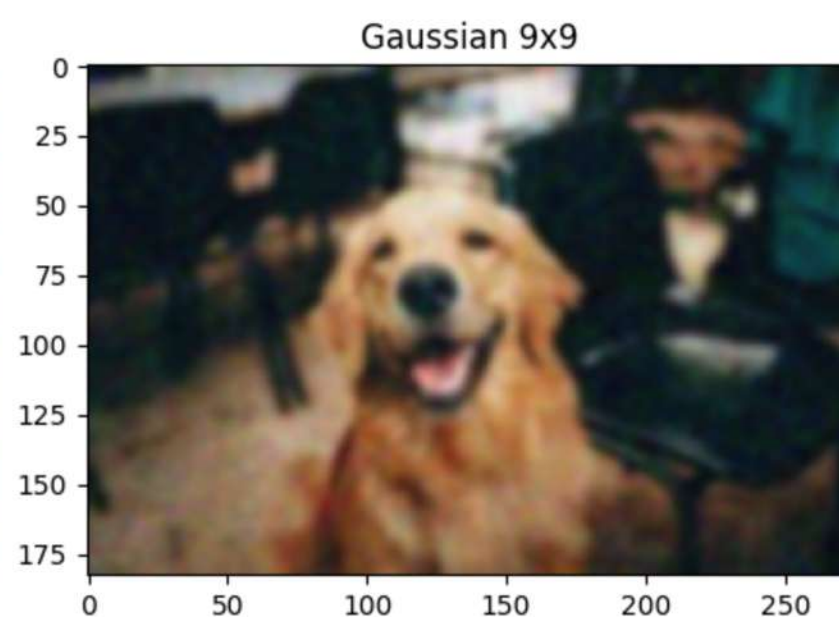
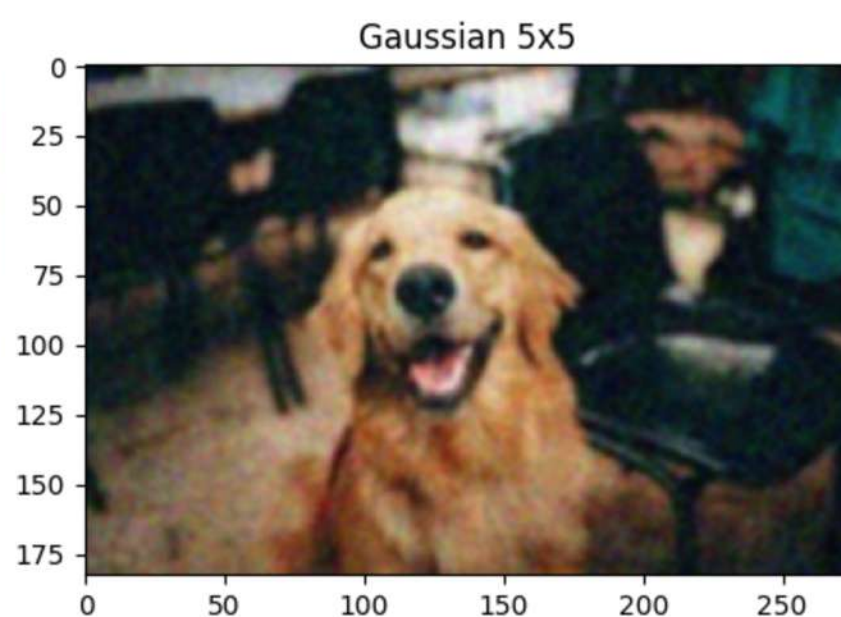
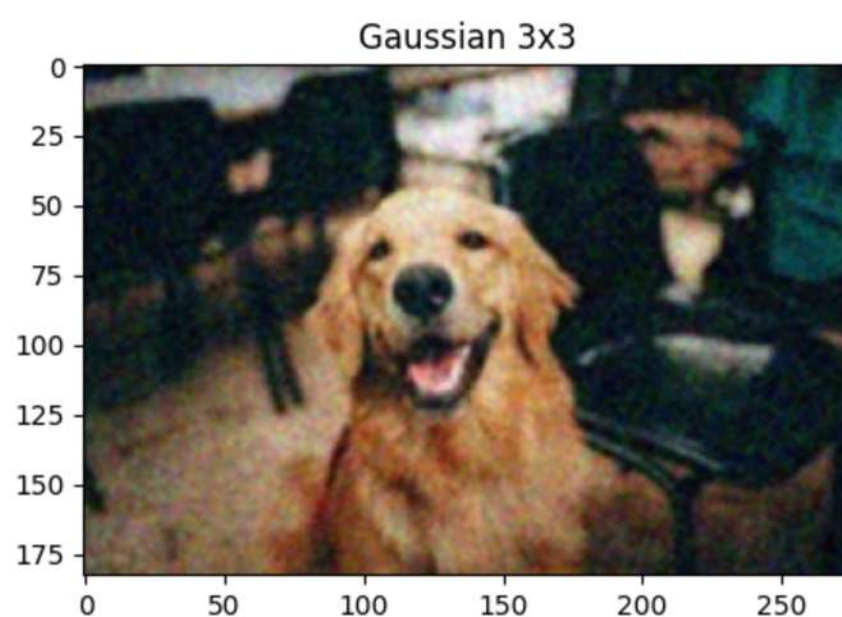
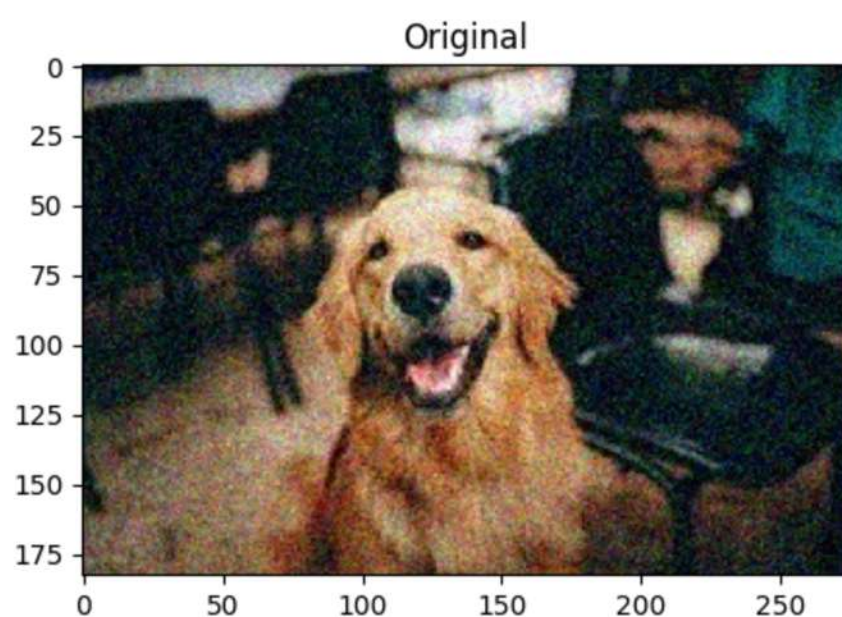
$$K_{3 \times 3} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$





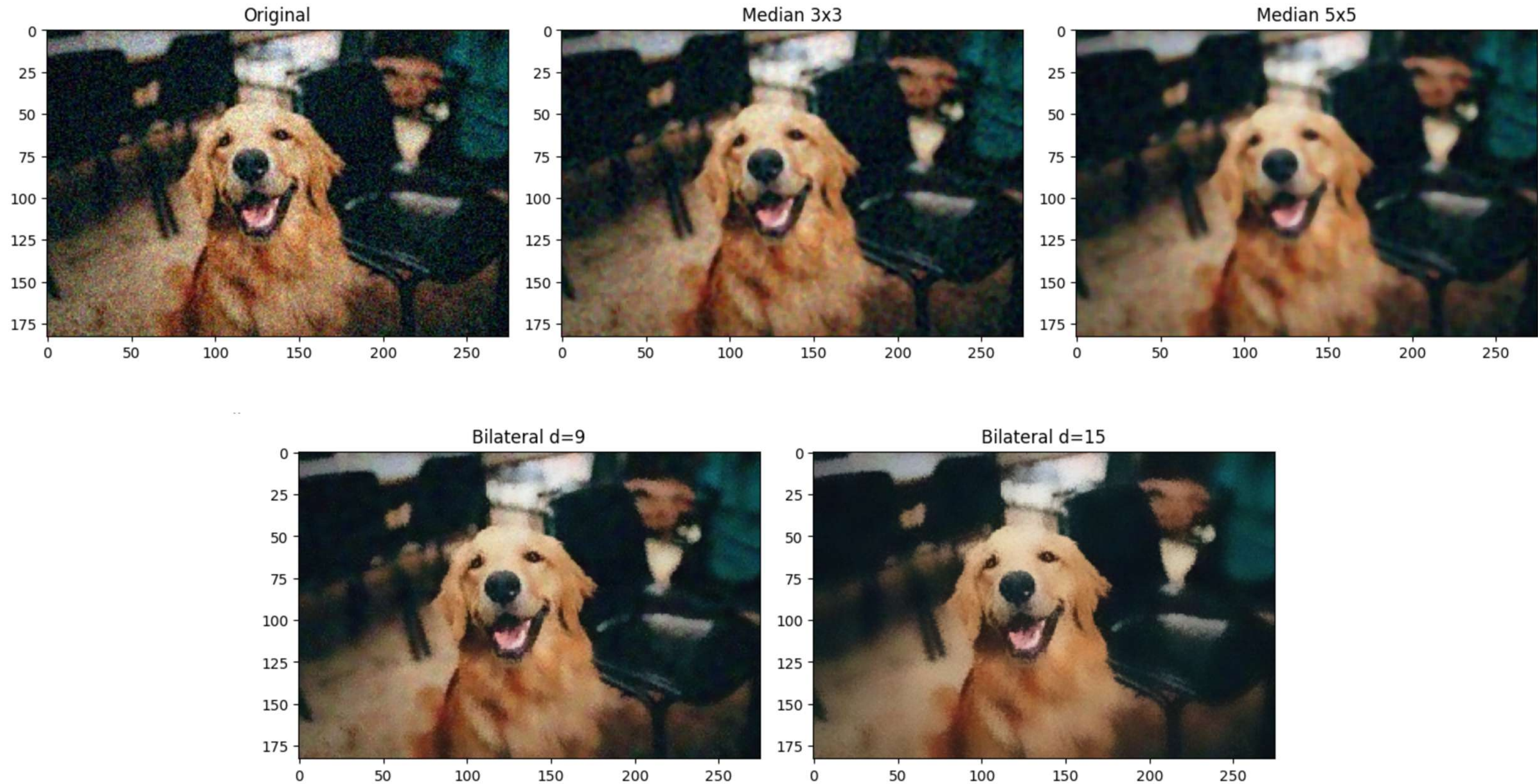
# Gaussian Filter

$$K_{3 \times 3} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$





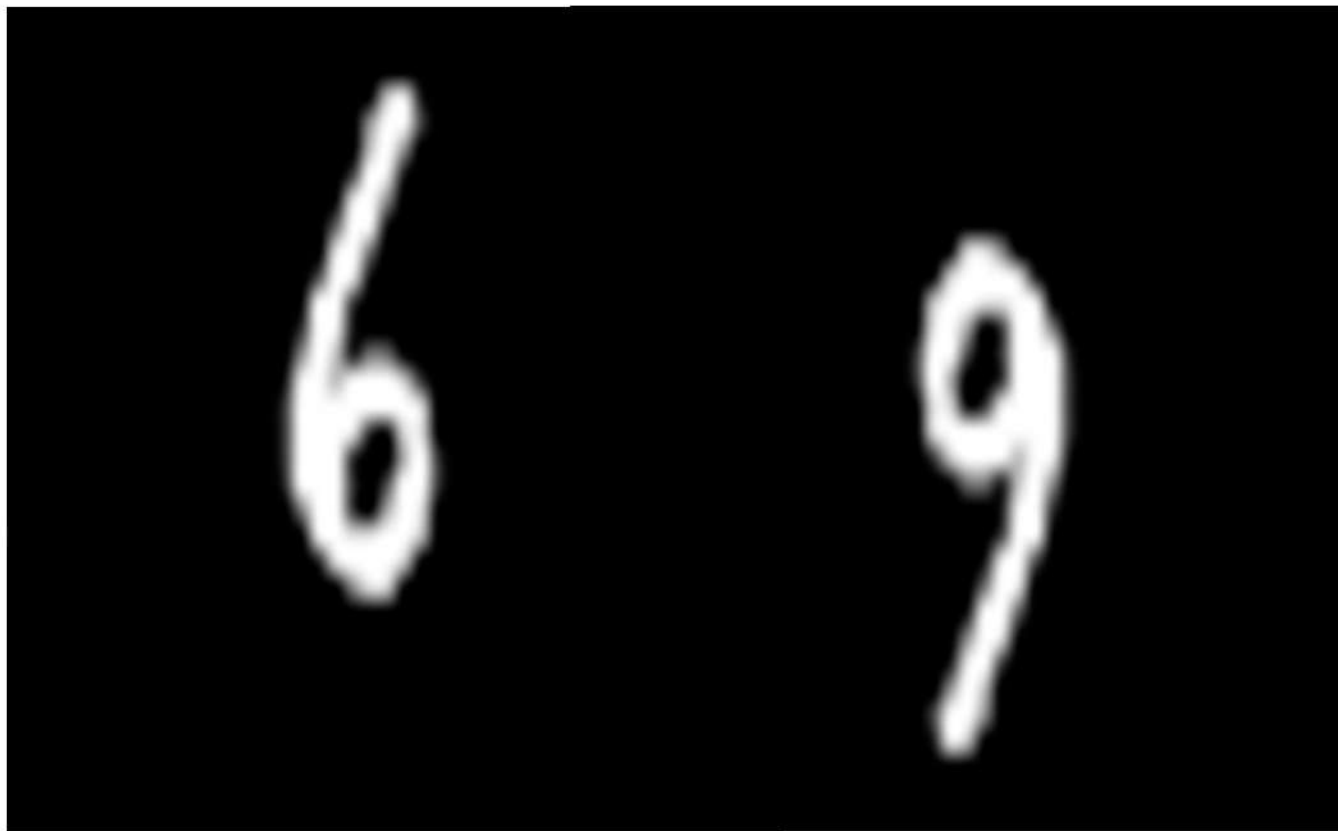
# Nonlinear Filter





# Data Augmentation

- Data Augmentation là tập hợp các kỹ thuật biến đổi dữ liệu gốc nhằm tạo ra thêm dữ liệu mới nhưng vẫn phải giữ nguyên ý nghĩa nhãn





- Data Augmentation được sử dụng trong giải đoạn chuẩn bị dữ liệu huấn luyện
  - Khi dữ liệu huấn luyện ít
  - Khi dữ liệu không đa dạng
  - Khi muốn tăng tính khái quát của mô hình
  - Khi có sự mất cân bằng giữa các lớp
- Việc thực hiện Data Augmentation chia làm hai kiểu:
  - Offline Augmentation
  - Online Augmentation



# Data Augmentation

rotate



shear



vertical-flip



horizontal-flip



crop



perspective-transform



sharpen



brighten



contrast



invert



gaussian-blur



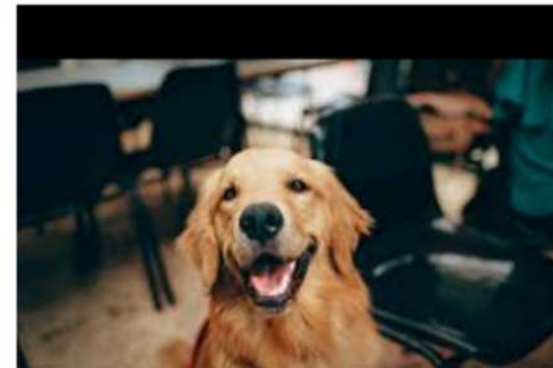
additive-noise



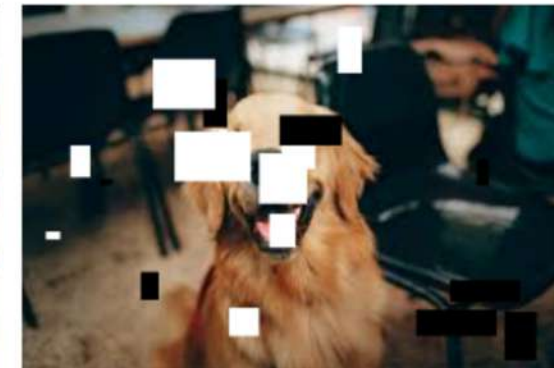
translate-x



translate-y



coarse-salt





# Ứng dụng trong bài toán thực tế

- Bộ dữ liệu: Clouds Photos
- Gồm 961 ảnh mây đã được gán nhãn, được chia thành 7 loại, gồm 3 tập train, validation và test



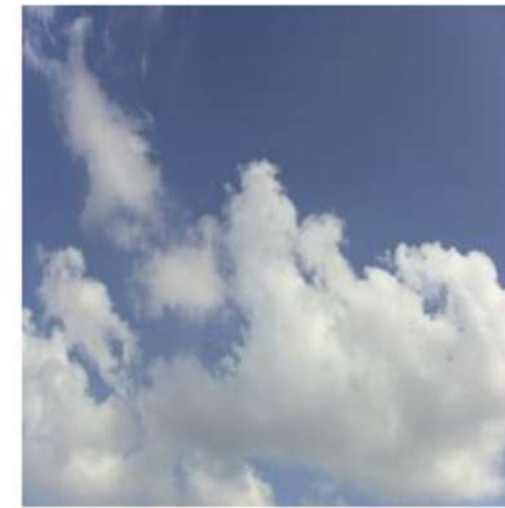
Cirriform



Clear Sky



Cumulonimbus



Cumulus



High Cumuliform

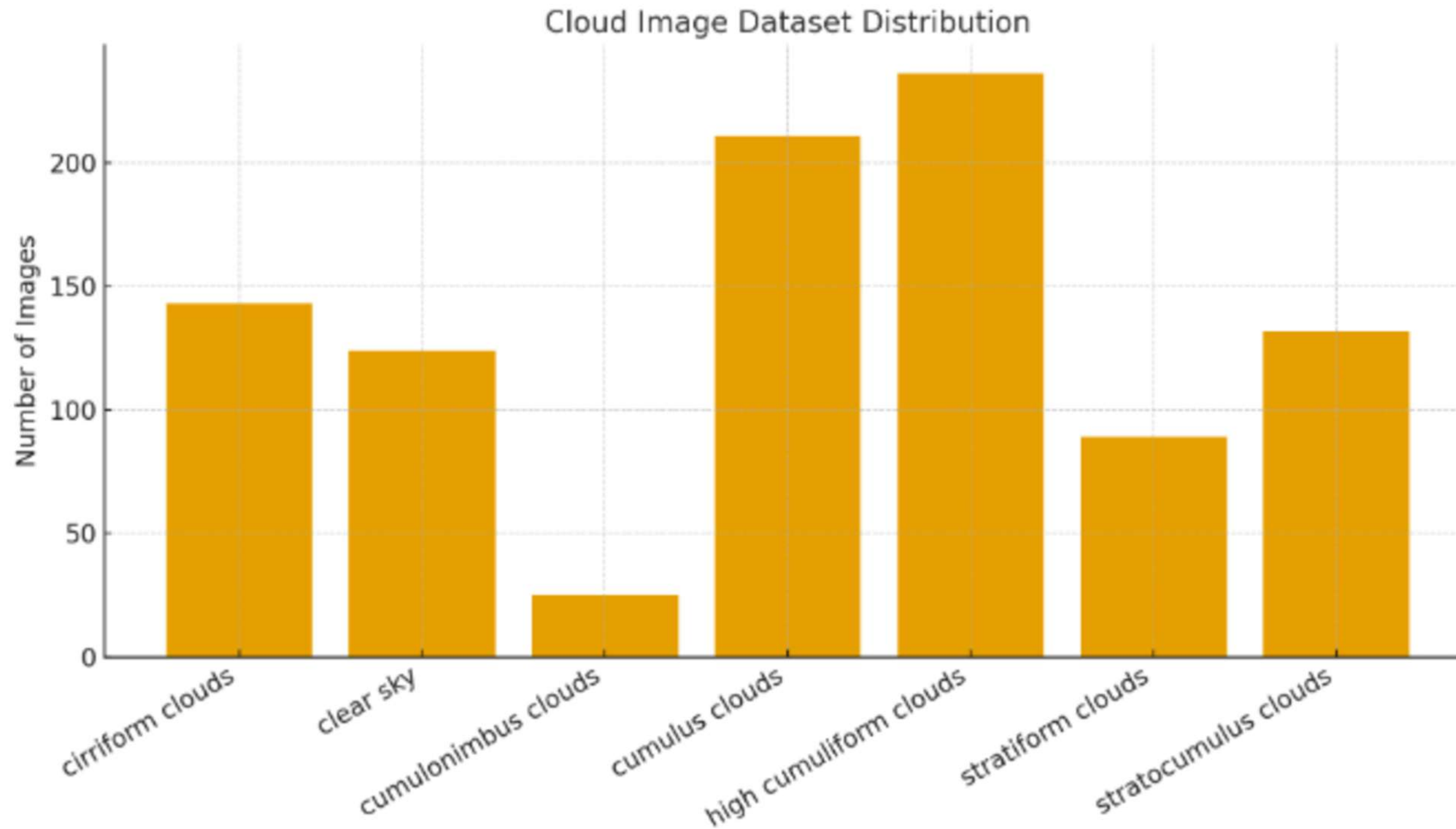


Stratiform



Stratocumulus

# Ứng dụng trong bài toán thực tế





MLP(

(net): Sequential(

(0): Flatten(start\_dim=1, end\_dim=-1)

(1): Linear(in\_features=256\*256\*3, out\_features=1024, bias=True)

(2): ReLU()

(3): Dropout(p=0.3, inplace=False)

(4): Linear(in\_features=1024, out\_features=512, bias=True)

(5): ReLU()

(6): Dropout(p=0.3, inplace=False)

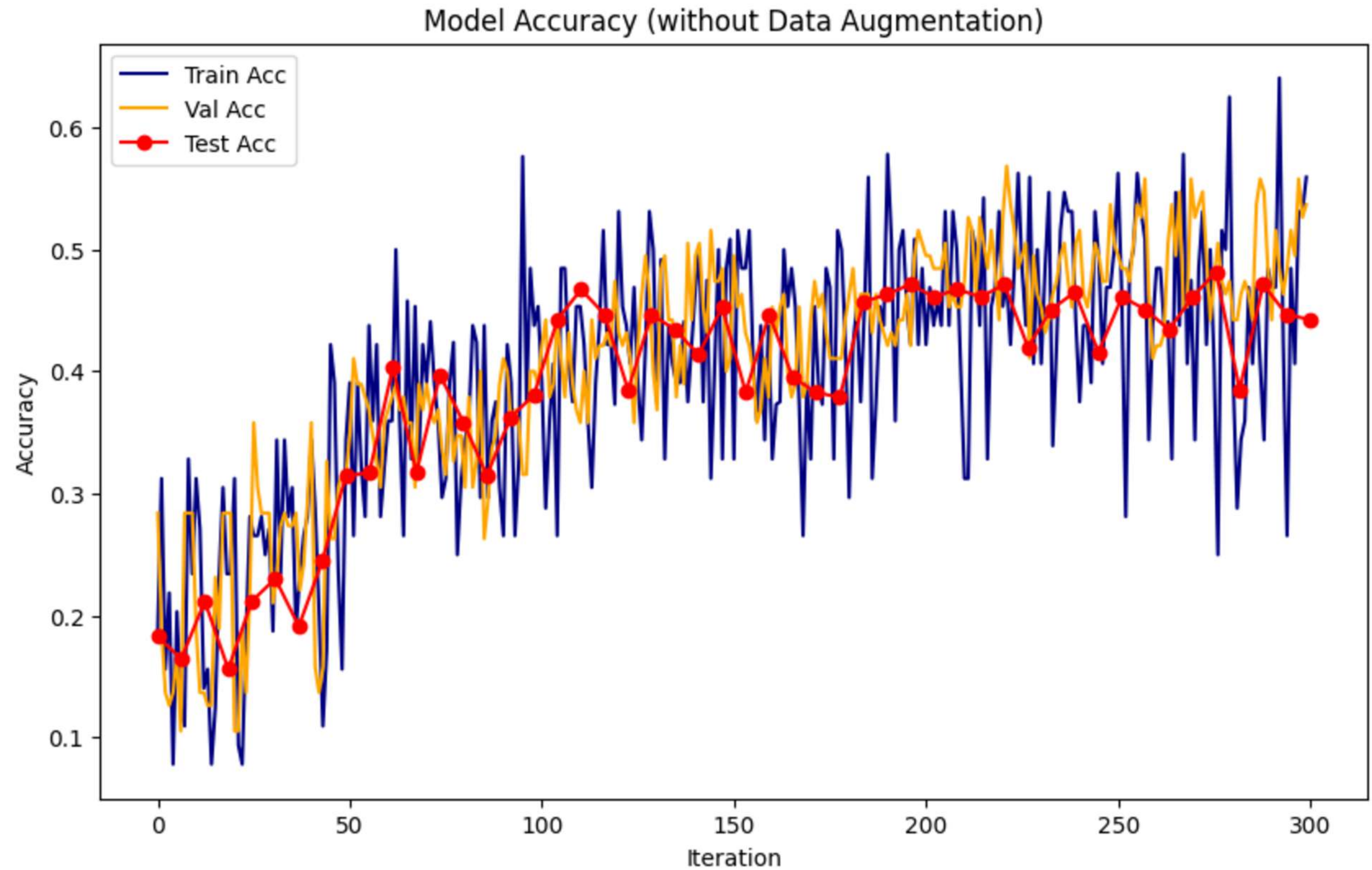
(7): Linear(in\_features=512, out\_features=7, bias=True)

)

)

# Ứng dụng trong bài toán thực tế

- Train Accuracy: 0.6815
- Validation Accuracy: 0.5766
- Test Accuracy: 0.4912

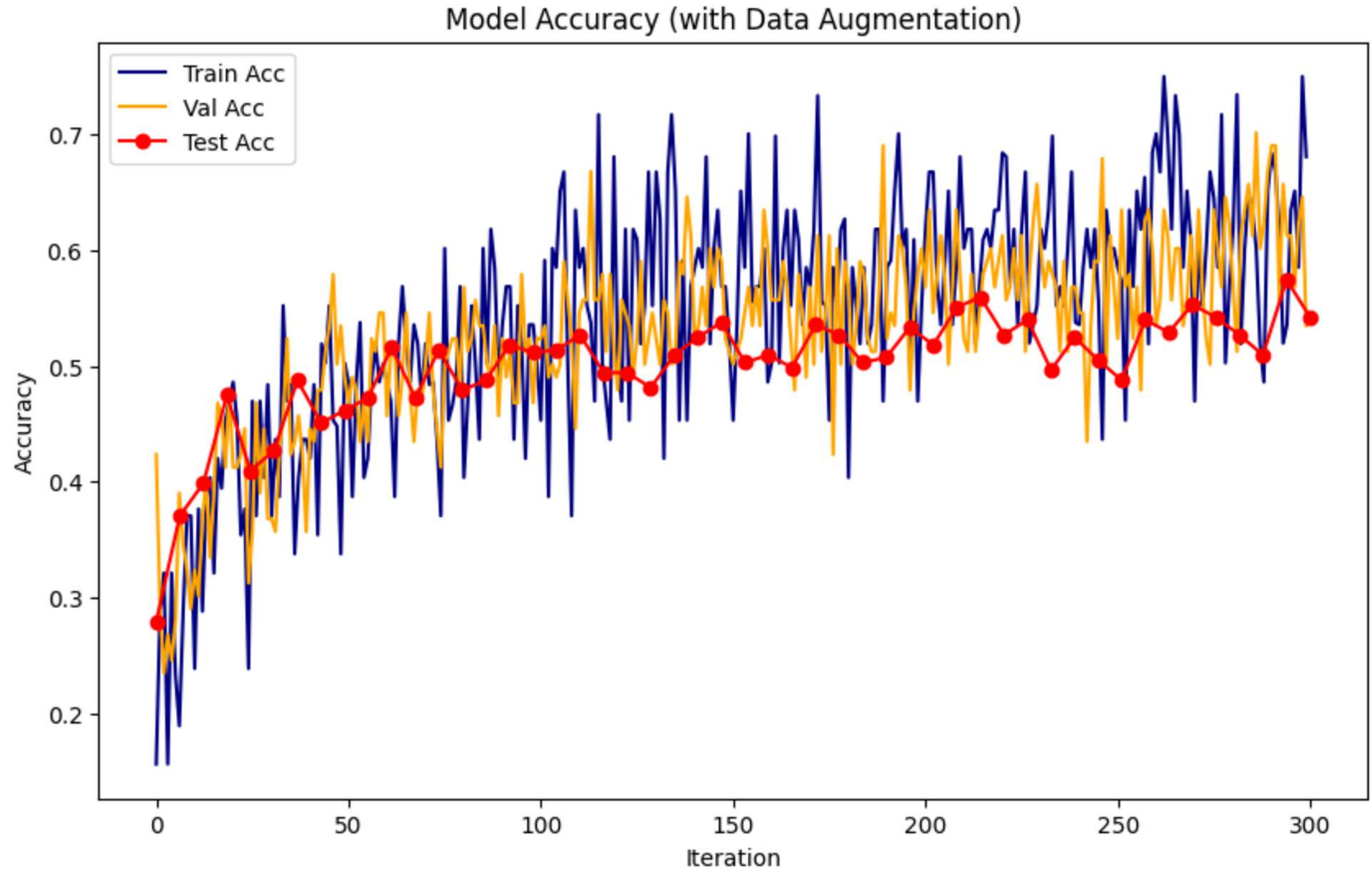




- Các phương pháp Augmentation thực hiện:
  - Standardize
  - Random Zoom In
  - Random Horizontal Flip
  - Random Rotation
  - Random Noise (Gaussian, SaltPepper)
  - Random Contrast

# Ứng dụng trong bài toán thực tế

- Train Accuracy: 0.7748
- Validation Accuracy: 0.6909
- Test Accuracy: 0.5788





# Ứng dụng trong bài toán thực tế

Model	Train Accuracy	Validation Accuracy	Test Accuracy	Runtime
Without Data Augmentation	0.6815	0.5766	0.4912	3m25s
With Data Augmentation	0.7748	0.6909	0.5788	11m5s

A large graphic on the left side of the slide. It features a dark blue background with a pattern of red dots of varying sizes arranged in concentric, slightly irregular circles, creating a sense of depth and movement. The word "HUST" is centered within this graphic in a white, bold, sans-serif font.

# HUST

# THANK YOU !



[hust.edu.vn](http://hust.edu.vn)



[fb.com/dhbkhn](https://fb.com/dhbkhn)