STUDIO
SHODWE

HOME SERVICE ABOUT US CONTACT US

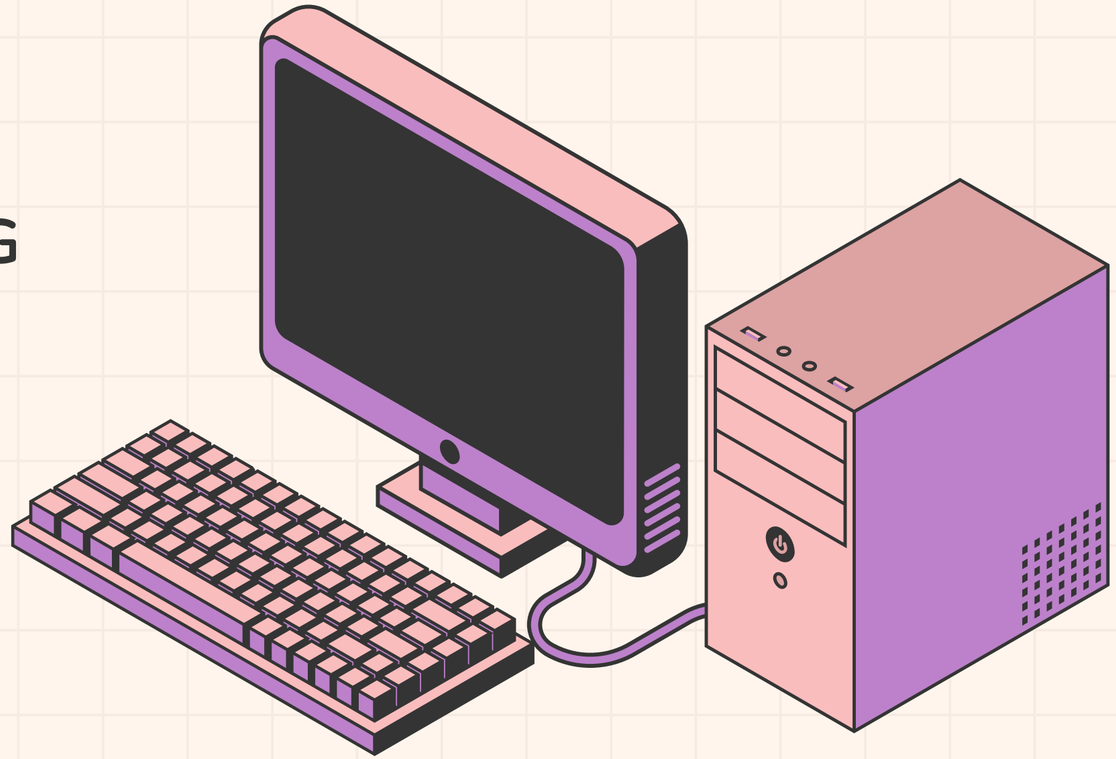
ĐỀ TÀI 6

NHẬN DẠNG CHỮ VIẾT VÀ HÌNH DẠNG ĐƠN GIẢN BẰNG MẠNG NEURAL

Nguyễn Nam Vũ - B22DCCN916

Nguyễn Quyết Tiến - B22DCCN724

Nguyễn Ngọc Thịnh - B22DCCN832



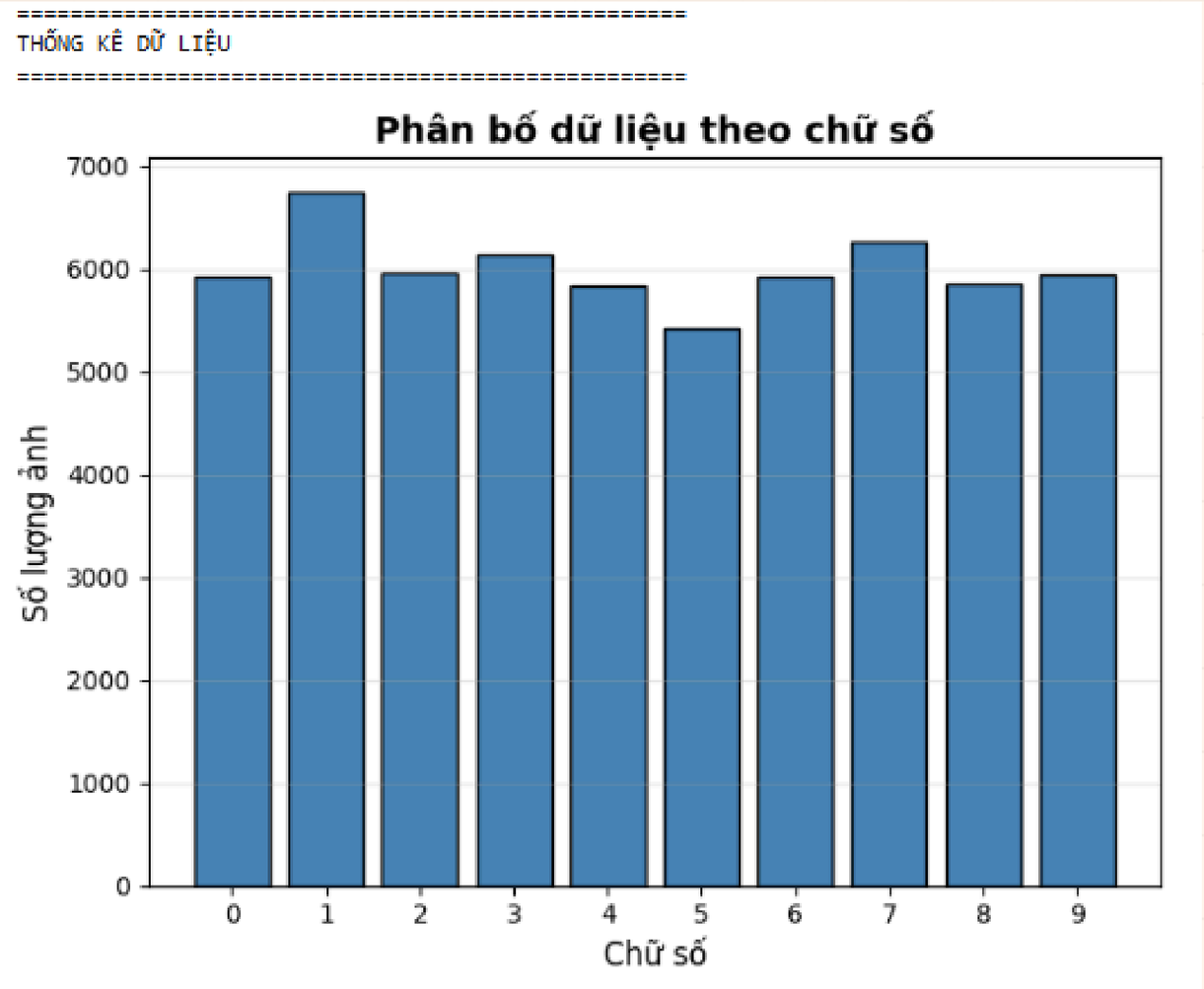
STUDIO
SHODWE

NHẬN DIỆN CHỮ SỐ VIẾT TAY

STUDIO
SHODWE

DATASET

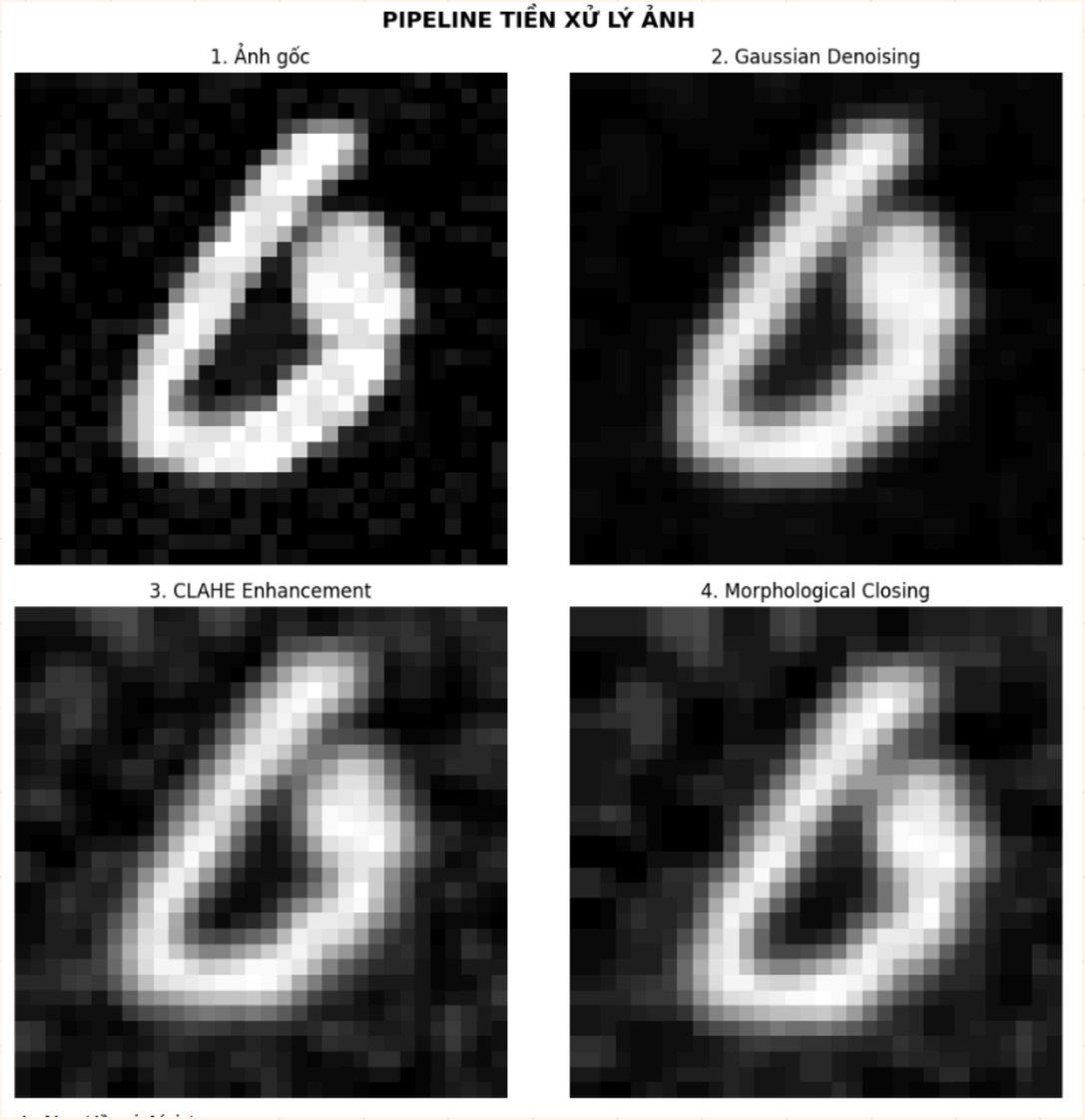
- Dataset: Corrupted MNIST (từ Kaggle).
- Số lượng mẫu: 60.000 hình ảnh chữ số viết tay.
- Định dạng: Grayscale, kích thước 28x28.



STUDIO
SHODWE

LÀM SẠCH ẢNH

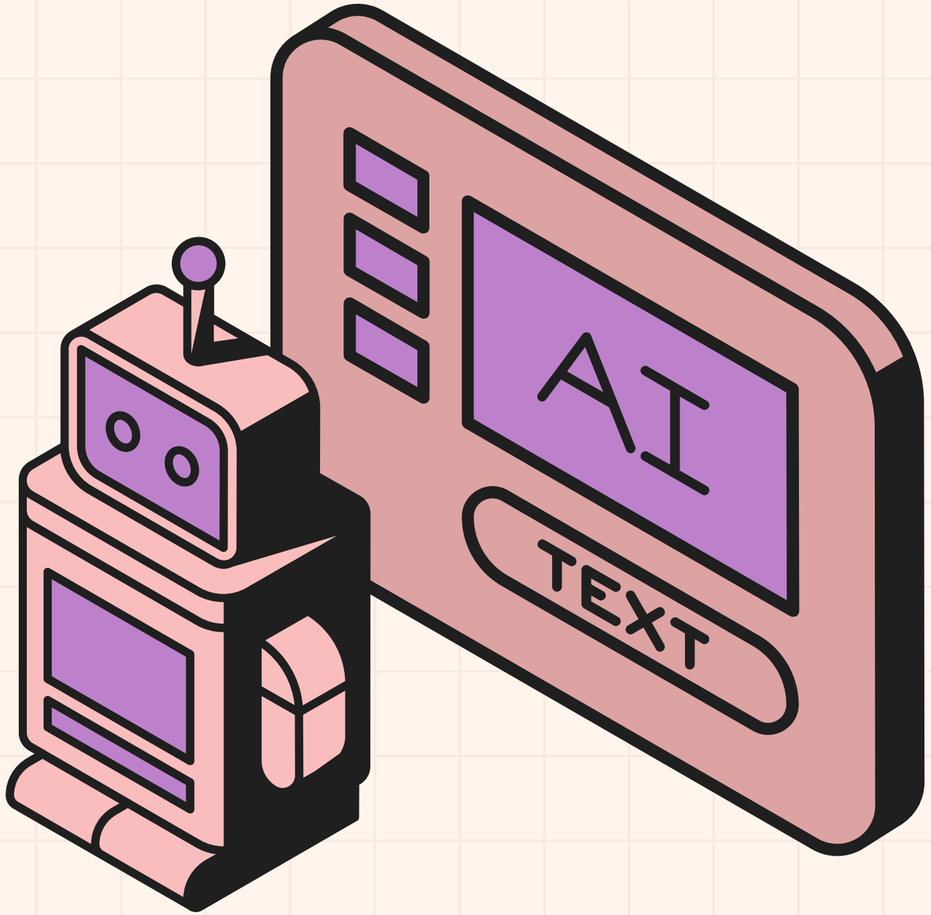
- Bước 1: Khử nhiễu, dùng cv2.GaussianBlur để làm mờ nhẹ các hạt sạn, nhiễu lấm tấm trên ảnh, giúp ảnh mịn hơn.
- Bước 2: Tăng độ tương phản: Sử dụng cv2.createCLAHE để làm cho nét chữ số đậm và nổi bật hẳn lên so với nền.
- Bước 3: Làm liền nét: Dùng cv2.morphologyEx để vá các vết đứt gãy nhỏ trên nét chữ, giúp chữ số trở nên liền mạch.



STUDIO
SHODWE

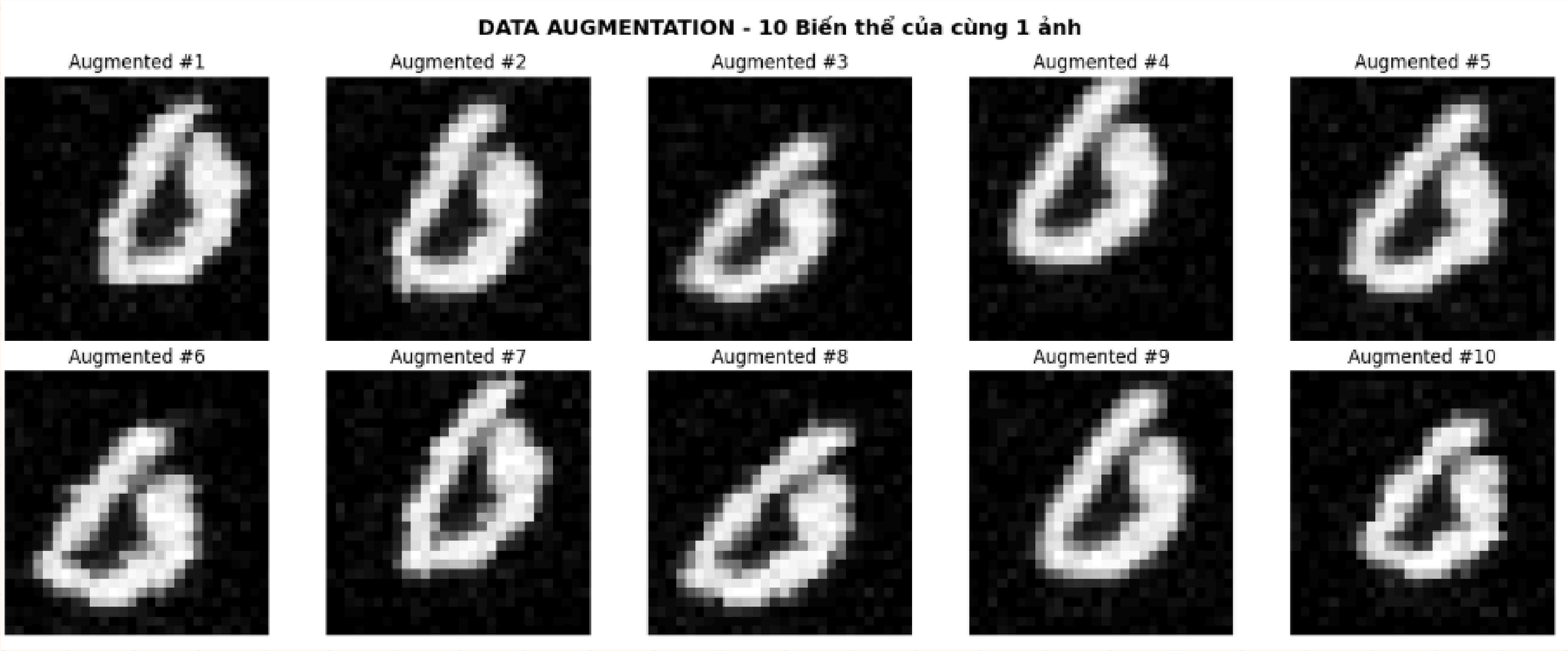
TĂNG CƯỜNG DỮ LIỆU THỜI GIAN THỰC

- Tại mỗi epoch, model sẽ nhìn 1 ảnh gốc với các góc nhìn khác nhau
- Các kỹ thuật tăng cường dữ liệu:
  - Xoay ảnh: Mỗi lần train, ảnh sẽ bị xoay ngẫu nhiên một góc từ -10 độ đến +10 độ.
  - Biến đổi hình học: Dịch chuyển sang trái/phải/lên/xuống (tối đa 10%), phóng to/ thu nhỏ (từ 90% đến 110%), éo xô lệch ảnh (tối đa 8 độ)
  - Biến đổi phối cảnh: Giả lập việc nhìn con số từ một góc nghiêng 3D (với độ méo là 15%, áp dụng với xác suất 25%)



STUDIO
SHODWE

TĂNG CƯỜNG DỮ LIỆU THỜI GIAN THỰC



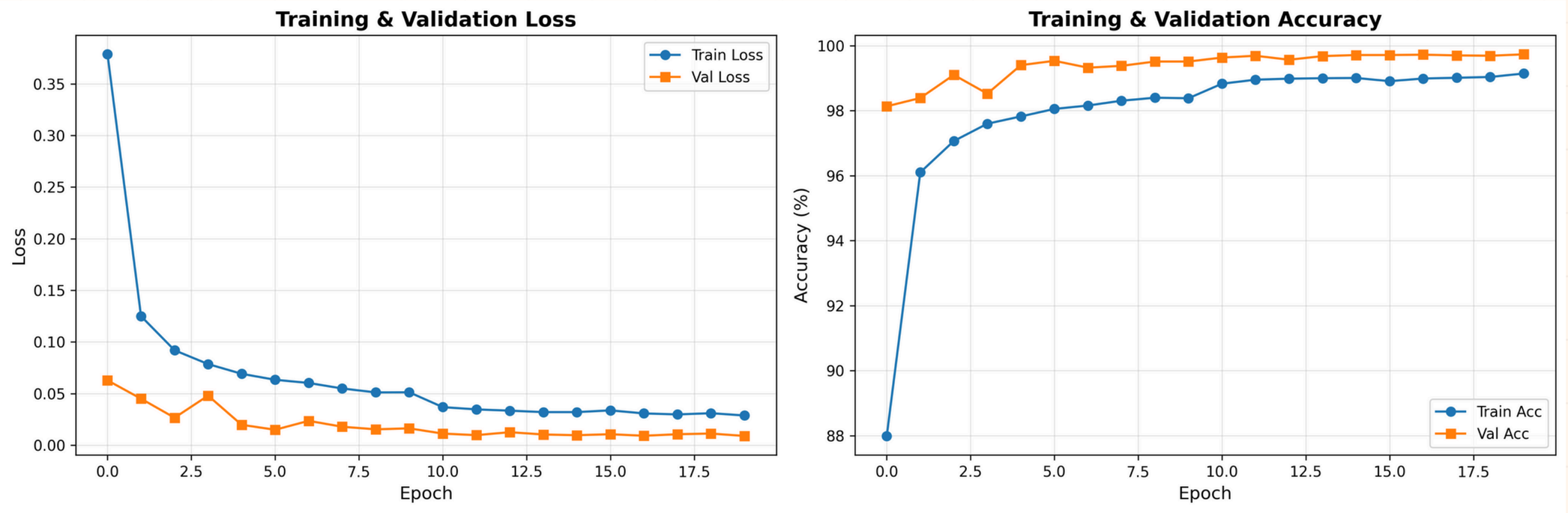
STUDIO SHODWE

CẤU HÌNH CNN MODEL

Layer	Input	Output	Tác dụng
Conv2d	[B, C_in, H, W]	[B, C_out, H, W]	Trích xuất đặc trưng (edges, patterns) qua kernel filters
BatchNorm2d	[B, C_in, H, W]	[B, C, H, W]	Chuẩn hóa dữ liệu, ổn định training, tăng tốc học
ReLU	[B, C_in, H, W]	[B, C, H, W]	Phi tuyến tính, giữ giá trị dương, loại âm (f(x)=max(0,x))
MaxPool2d	[B, C_in, H, W]	[B, C, H/2, W/2]	Giảm kích thước, giữ đặc trưng quan trọng nhất
Dropout	[B, N]	[B, N]	Tắt ngẫu nhiên neurons, chống overfitting khi train

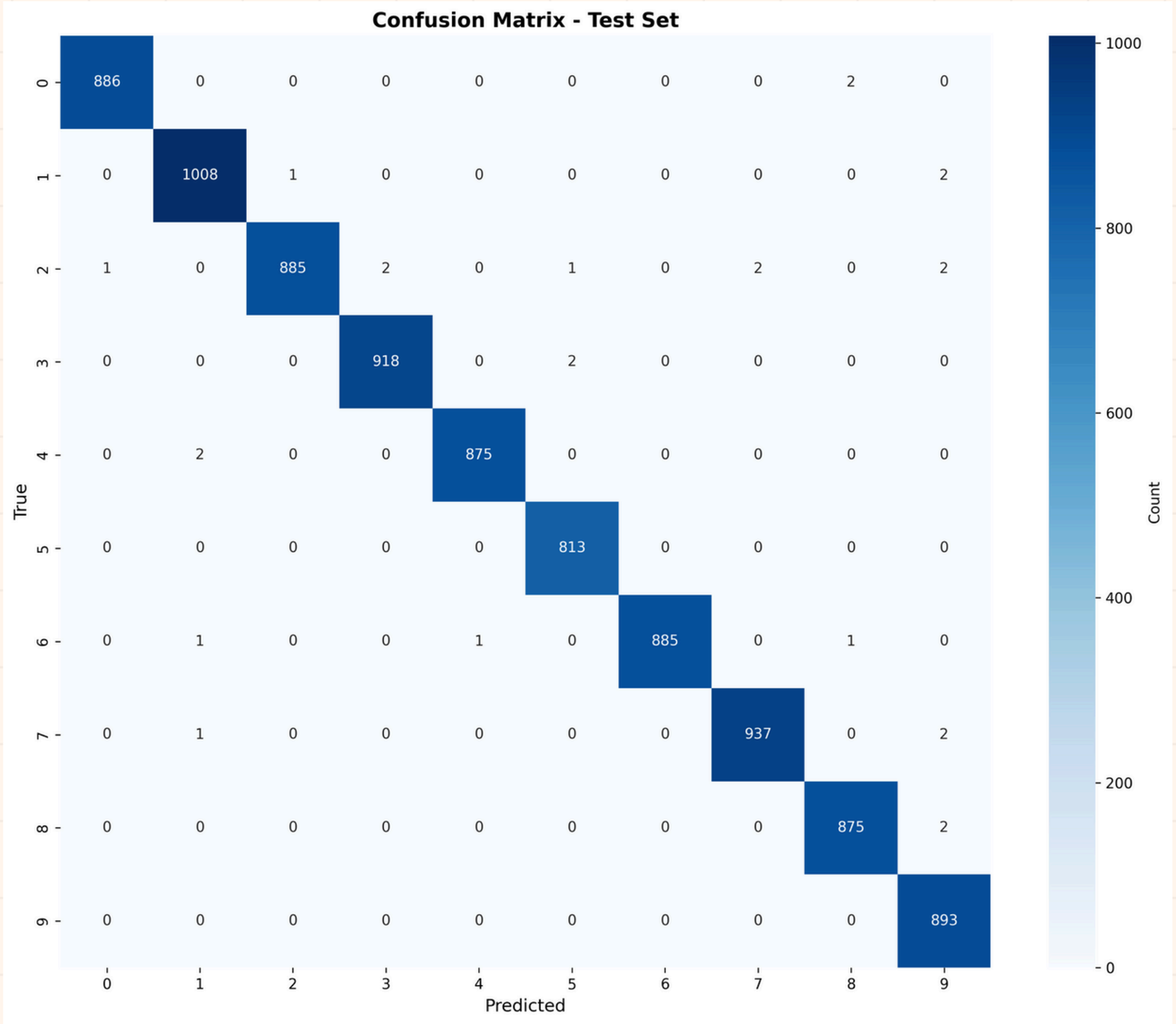
STUDIO
SHODWE

KẾT QUẢ HUẤN LUYỆN



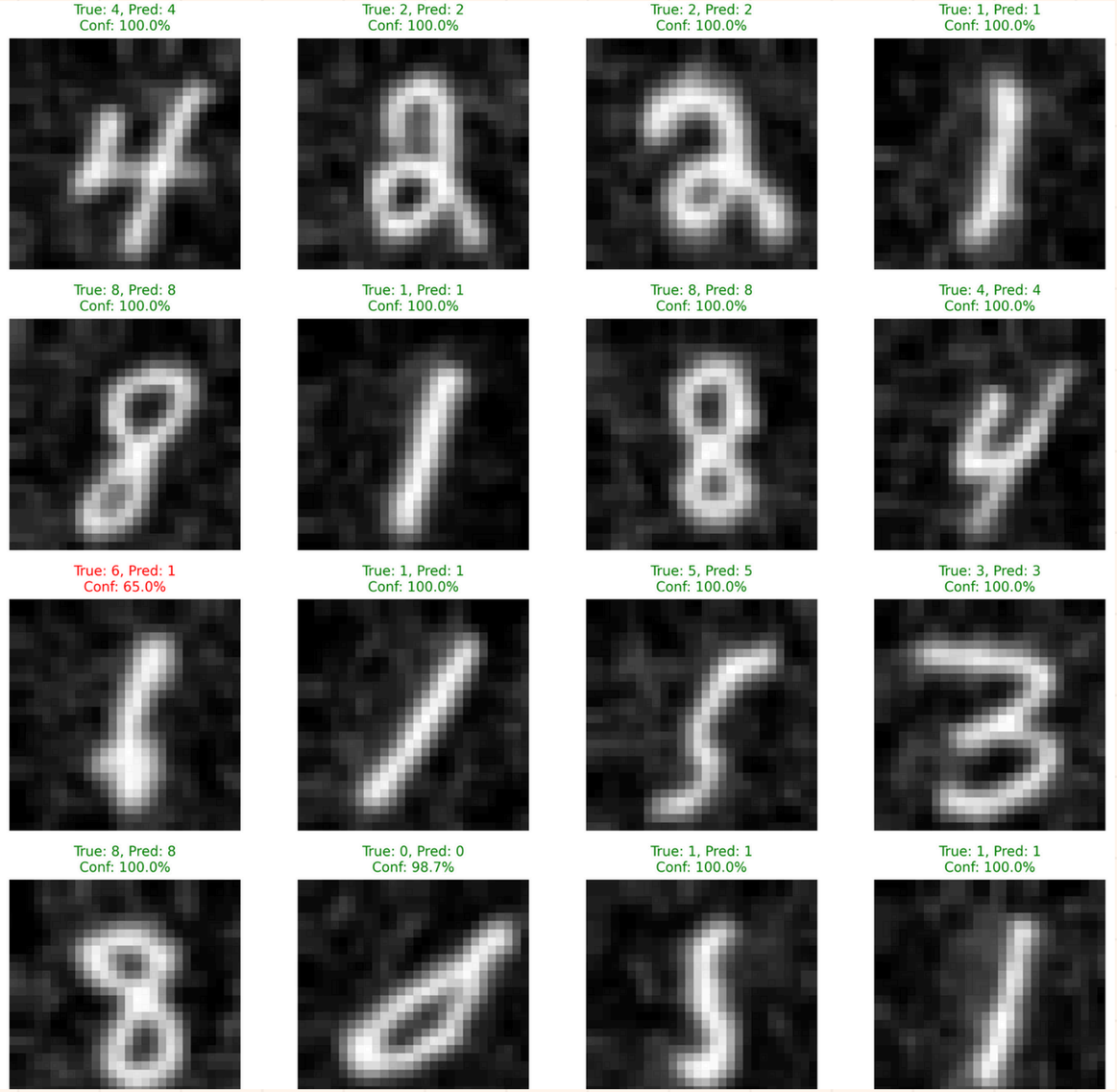
STUDIO
SHODWE

THỬ NGHIỆM TRÊN TẬP TEST



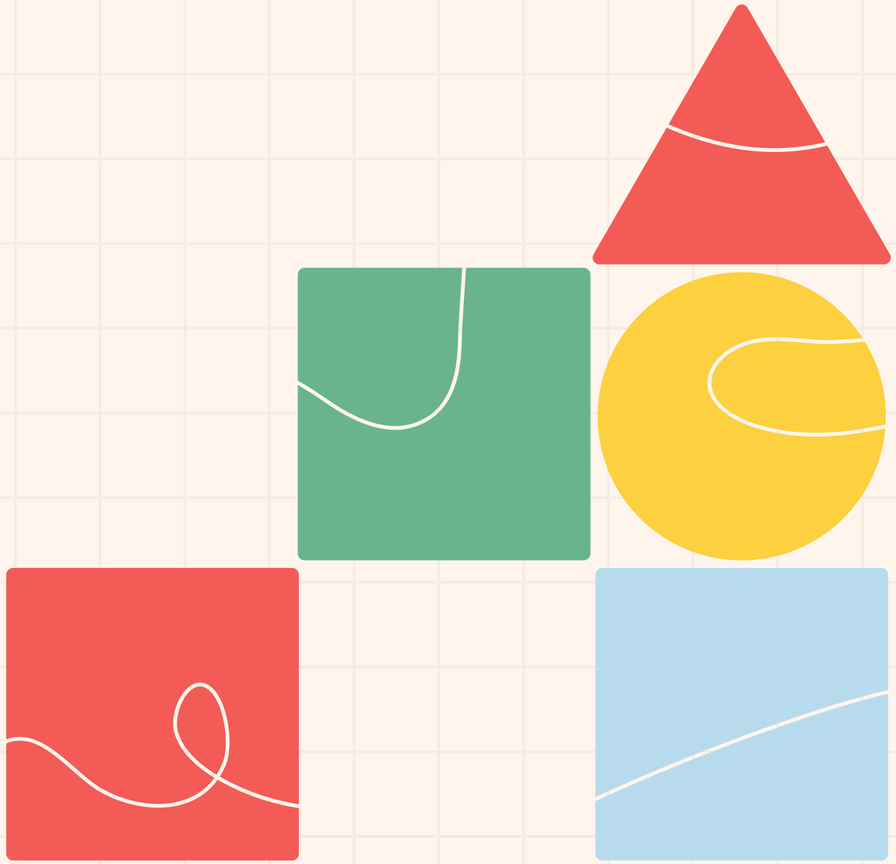
STUDIO
SHODWE

THỬ NGHIỆM TRÊN TẬP TEST



STUDIO
SHODWE

PHÁT HIỆN HÌNH HỌC CƠ BẢN



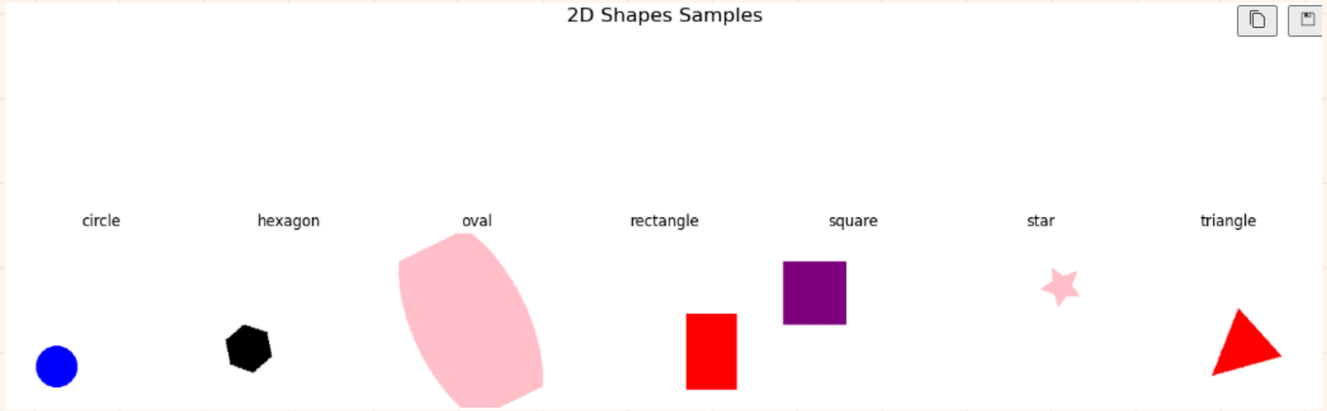
STUDIO SHODWE

DATASET

- Dataset: 2D Geometric Shapes Dataset (từ Kaggle).
- Sử dụng 7 hình học cơ bản, mỗi hình có 50000 ảnh
- Kích thước: 224x224 pixels

```
Total shapes samples: 350000

Shapes distribution:
label
circle       50000
hexagon      50000
oval         50000
rectangle    50000
square       50000
star         50000
triangle     50000
Name: count, dtype: int64
```



STUDIO SHODWE

TĂNG CƯỜNG DỮ LIỆU

- Resize(64,64): Chuyển mọi ảnh về kích thước 64x64
- RandomRotation(15): Xoay ảnh ngẫu nhiên trong khoảng +- 15 độ
- RandomHorizontalFlip(): Lật ảnh với xác suất 50%
- ToTensor()
  - Chuyển ảnh từ PIL hoặc NumPy → Tensor PyTorch
  - Đưa giá trị pixel từ 0-255 về 0-1
- Normalize: Chuẩn hóa giá trị từng kênh ảnh(R,G,B)

```
shapes_train_transform = transforms.Compose([
    transforms.Resize((64, 64)),
    transforms.RandomRotation(15),
    transforms.RandomHorizontalFlip(),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])
])

shapes_test_transform = transforms.Compose([
    transforms.Resize((64, 64)),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])
])
```

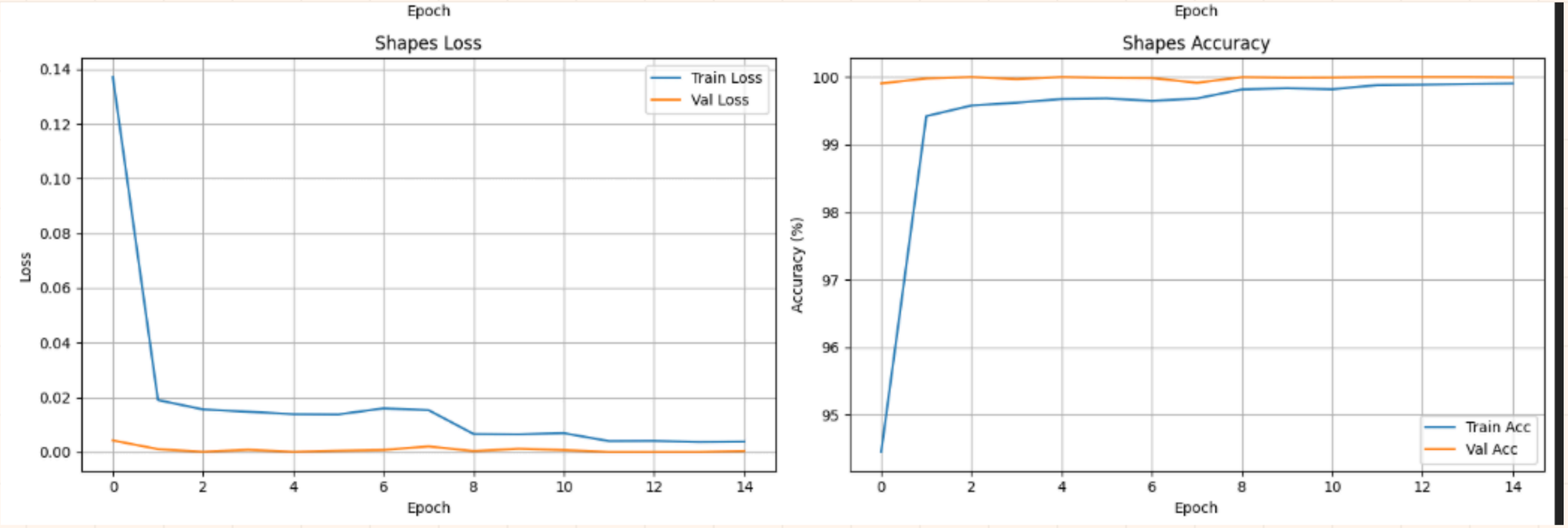
STUDIO SHODWE

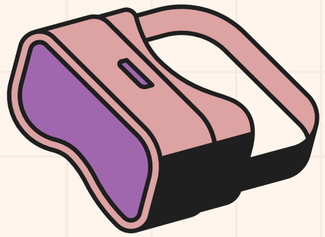
CẤU HÌNH CNN MODEL

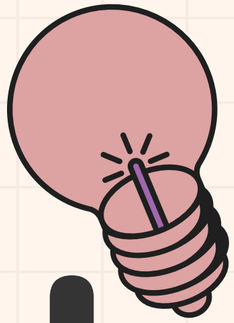
Thành phần	Thông số	Kích thước đầu ra	Chức năng
Input Image	3 × 64 × 64	3 × 64 × 64	Ảnh RGB đầu vào
Conv1	Conv2d(3 → 32, kernel=3, padding=1) + ReLU	32 × 64 × 64	Trích đặc trưng cạnh và góc
MaxPool1	2×2	32 × 32 × 32	Giảm kích thước, giảm nhiễu
Conv2	Conv2d(32 → 64, kernel=3, padding=1) + ReLU	64 × 32 × 32	Học đặc trưng phức tạp hơn
MaxPool2	2×2	64 × 16 × 16	Giảm kích thước
Conv3	Conv2d(64 → 128, kernel=3, padding=1) + ReLU	128 × 16 × 16	Nhận diện hình dạng rõ hơn
MaxPool3	2×2	128 × 8 × 8	Giảm kích thước
Conv4	Conv2d(128 → 256, kernel=3, padding=1) + ReLU	256 × 8 × 8	Trích đặc trưng sâu
MaxPool4	2×2	256 × 4 × 4	Output để đưa vào FC
Flatten	-	256 × 4 × 4 = 4096	Chuyển thành vector
FC1	Linear(4096 → 512) + ReLU	512	Học phân bố đặc trưng
Dropout	p = 0.5	512	Giảm overfitting
FC2 (Output)	Linear(512 → 7)	7 lớp	Dự đoán 7 loại hình học

STUDIO
SHODWE

KẾT QUẢ HUẤN LUYỆN







THANK YOU

+123-456-7890 WWW.REALLYGREATSITE.COM 123 ANYWHERE ST., ANY CITY