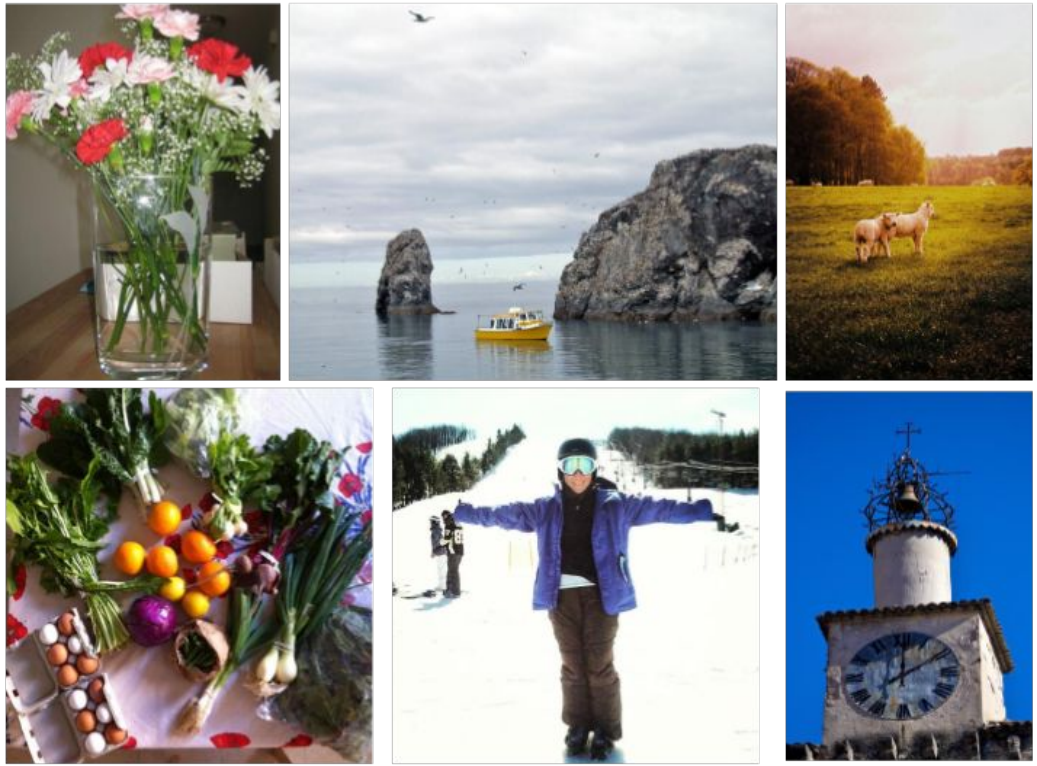
Automated Image Colorization

Revathi Bhuvaneswari Tianying Luo Yue Zeng

Dataset

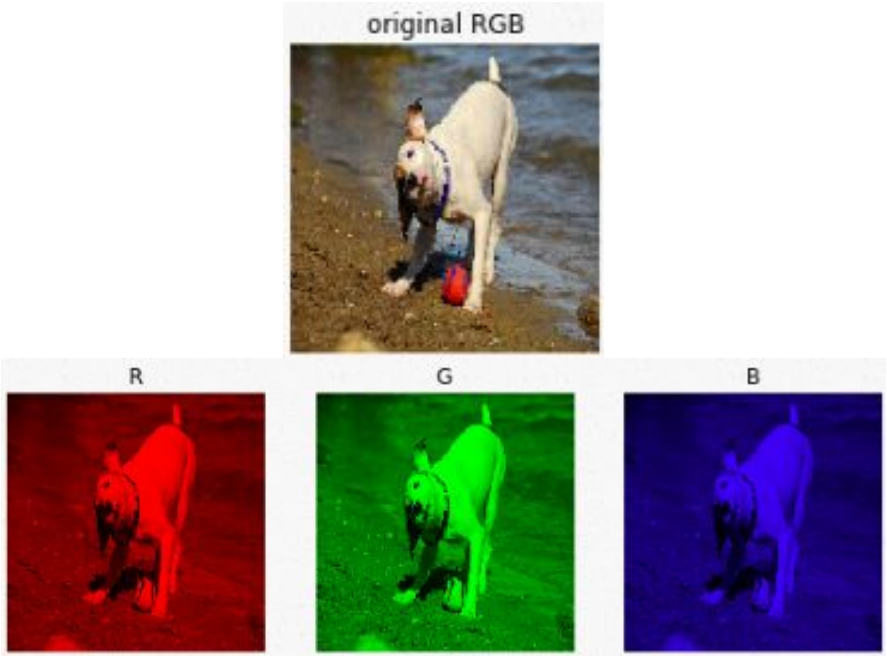


16, 590 images

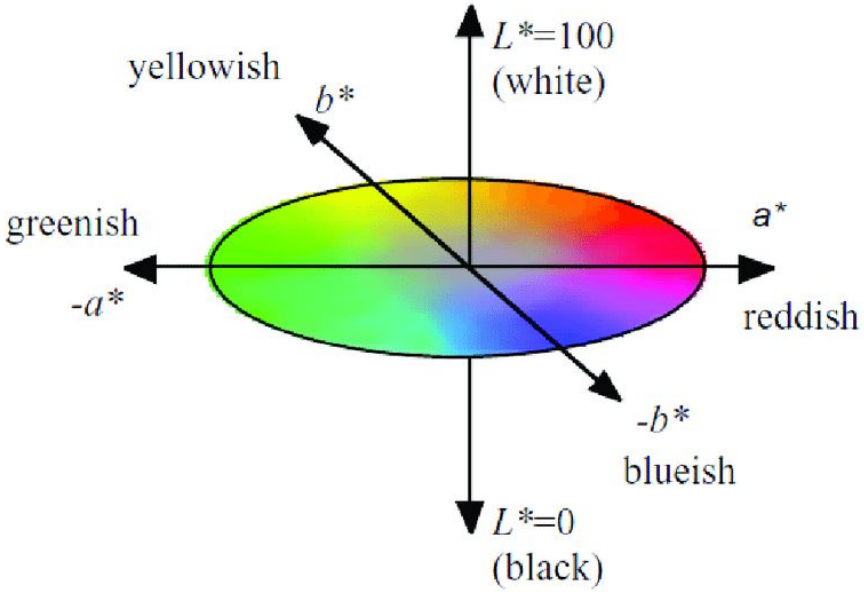
TRAIN	70 %
VAL	20 %
TEST	10 %

Colorspace

RGB



LAB



Evaluation Metrics

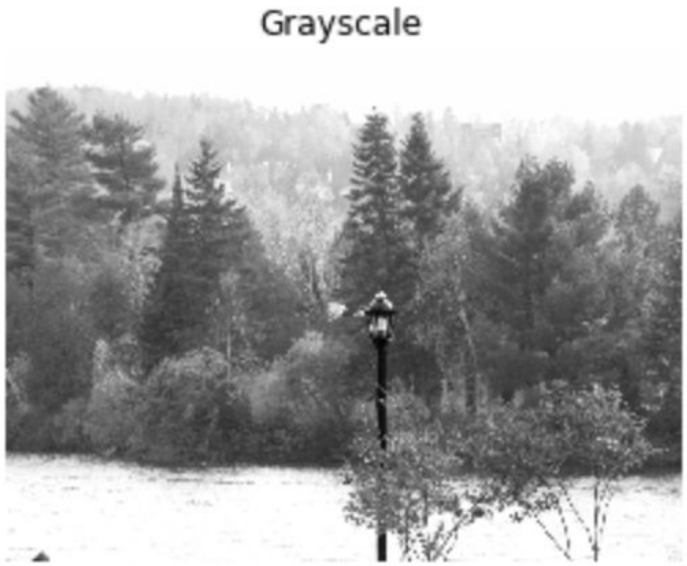
L1 (OR) MSE LOSS

$$\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

L2 (OR) MSE LOSS

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Baseline CNN



Left - Original(output)

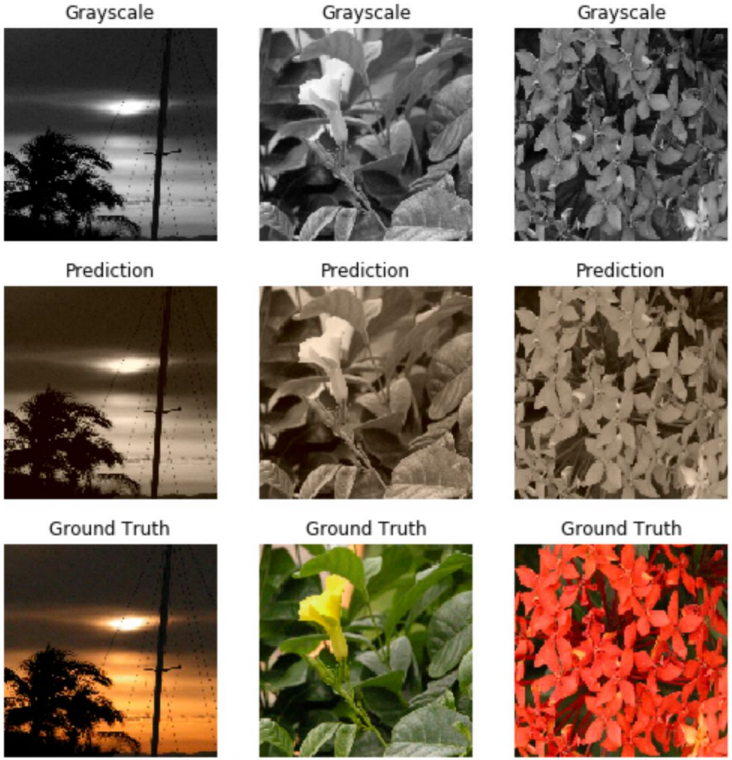
Right - Grayscale(input)

Baseline CNN

Input: 224 x 224 x 1	
3 x 3 Conv, 64, S = 2	224x 224x 64
3 x 3 Conv, 64, S = 2	112x 112 x 64
3 x 3 Conv, 128, S = 2	112x112 x128
3 x 3 Conv, 128, S = 2	56 x 56 x128
3 x 3 Conv, 256, S = 2	56 x 56 x256
3 x 3 Conv, 256, S = 2	28 x 28 x 256
3 x 3 Conv, 512, S = 2	28 x28x1512
3 x 3 Conv, 512, S = 2	14x214x 512
3 x 3 Conv, 512, S = 2	28 x28x512
3 x 3 Conv, 256, S = 2	56 x56x256
3 x 3 Conv, 128, S = 2	112x112x 128
3 x 3 Conv, 64, S = 2	224 x224x 64
3 x 3 Conv, 256, S = 2	224 x224x2
Output: 224 x 224 x 2	

- L2 Loss
- ReLU / TanH
- Batch Normalization
- No Dropout
- He Normal Weights
- Transpose Layers

Results (Failure)



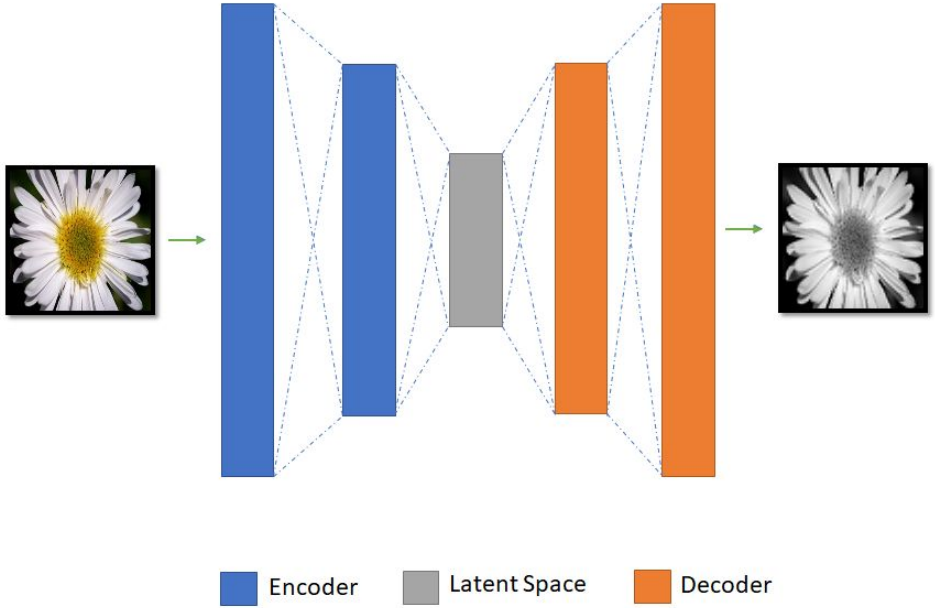
- L2 Loss
- ReLU / TanH
- Batch Normalization
- Dropout
- He Normal Weights

Results(Success)

Epochs = 500

Epochs = 1000

Baseline Autoencoder



Baseline Autoencoder

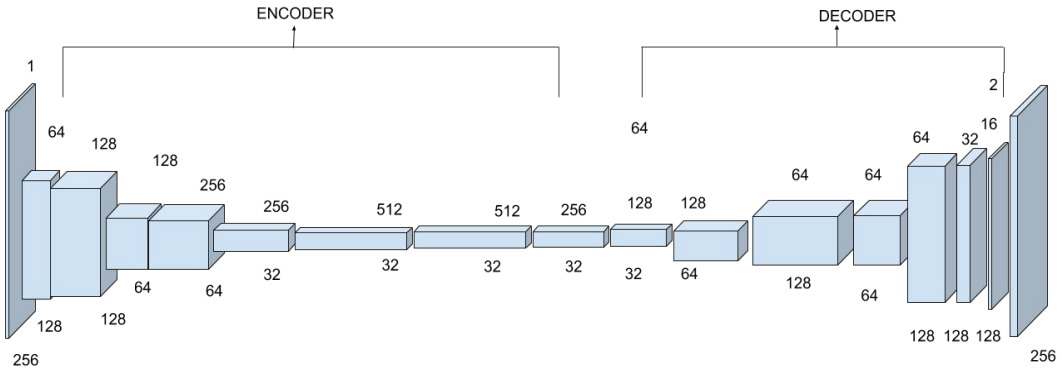
1. **Colorization Autoencoder using RGB**

Input: 224 x 224 x 1	
3 x 3 Conv, 64, S = 2	112 x112 x 64
3 x 3 Conv, 128, S = 2	56 x 56 x 128
3 x 3 Conv, 256, S = 2	28 x 28 x 256
3 x 3 Conv, 512, S = 2	14 x 14 x512
Flatten	FC: 100352
Dense, 512	FC: 512

↓

Dense, 512	FC: 512
Reshape	FC: 100352
3 x 3 Conv, 512, S = 2	28 x 28 x512
3 x 3 Conv, 256, S = 2	56 x 56 x 256
3 x 3 Conv, 128, S = 2	112 x112x128
3 x 3 Conv, 64, S = 2	224 x224x 64
Input: 224 x 224 x 3	

2. **Colorization Autoencoder using LAB**

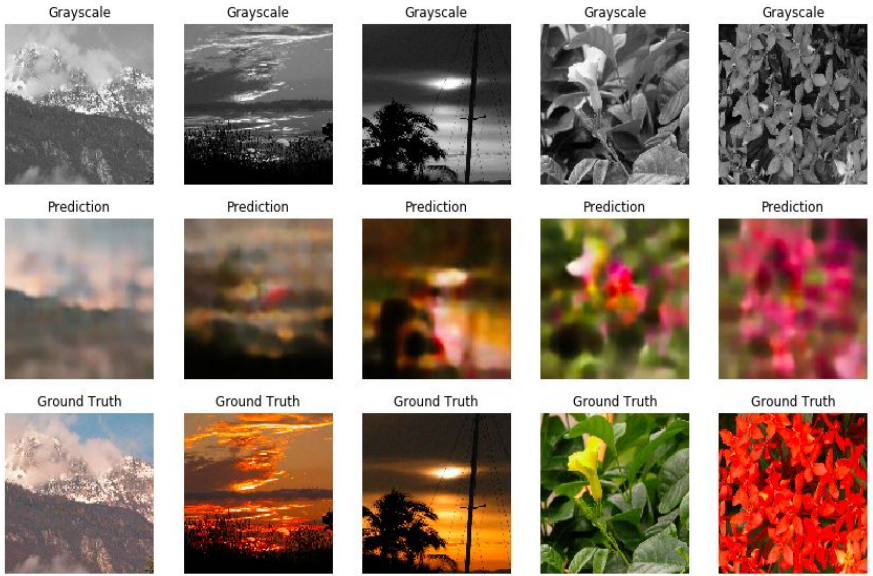


Baseline Autoencoder

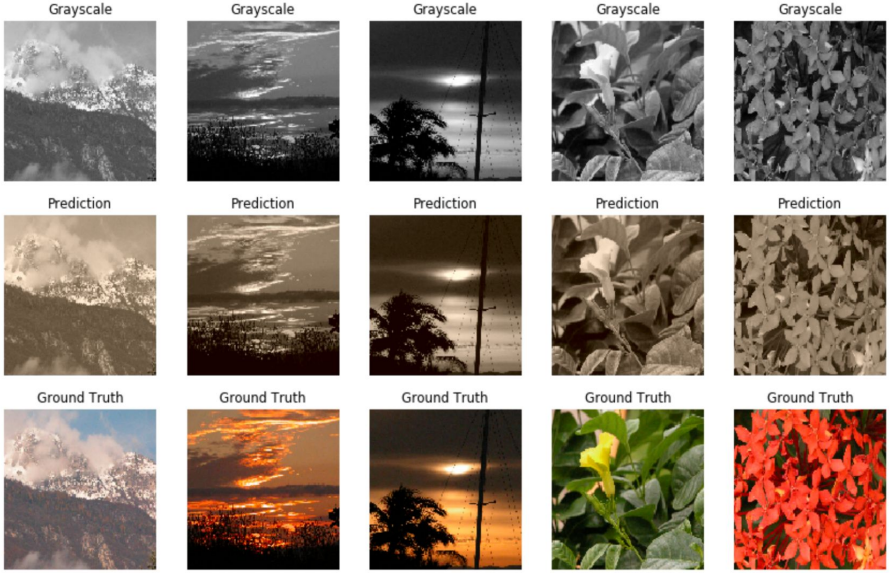
- **L2 Loss**
- **ReLU / TanH**
- **Batch Normalization**
- **Transpose Layers**
- **Upsampling Layers**
- **Callbacks=[ReduceLROnPlateau, ModelCheckpoint]**

Results

1. Colorization Autoencoder using RGB

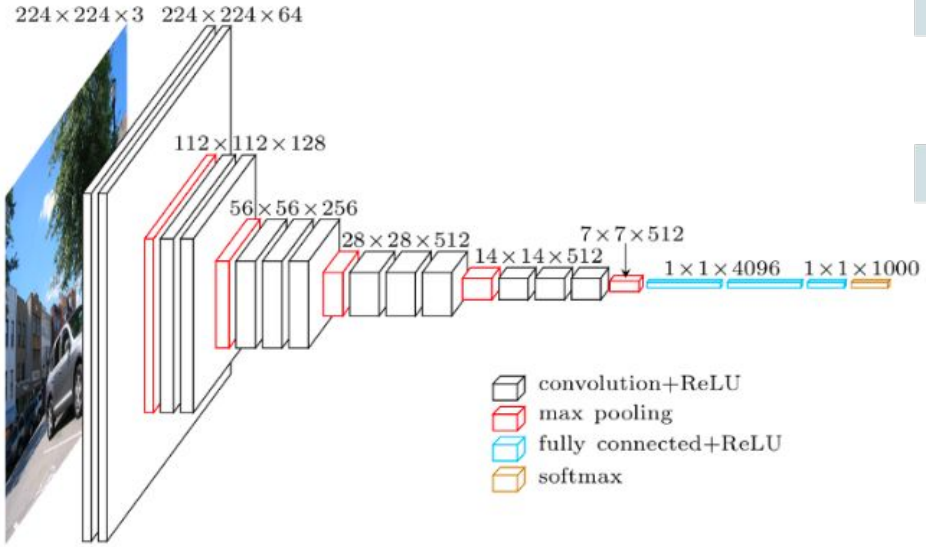


2. Colorization Autoencoder using LAB



Transfer Learning: VGG-16

1. VGG16 Model Architecture



2. Custom Decoder Layers

ENCODER	Input: 7 x 7 x 512	
	Output: 7 x 7 x 512	
DECODER	Input: 7 x 7 x 512	
	3 x 3 Conv, 256	7 x 7 x 256
	3 x 3 Conv, 128	7 x 7 x 128
	2 x 2, Upsampling	14 x 14 x 128
	3 x 3 Conv, 64	14 x 14 x 64
	2 x 2, Upsampling	28 x 28 x 64
	3 x 3 Conv, 32	28 x 28 x 32
	2 x 2, Upsampling	56 x 56 x 32
	3 x 3 Conv, 64, S = 1	56 x 56 x 16
	2 x 2, Upsampling	112 x 112 x 16
	3 x 3 Conv, 32, S = 1	112 x 112 x 2
	2 x 2, Upsampling	224 x 224 x 2

Results

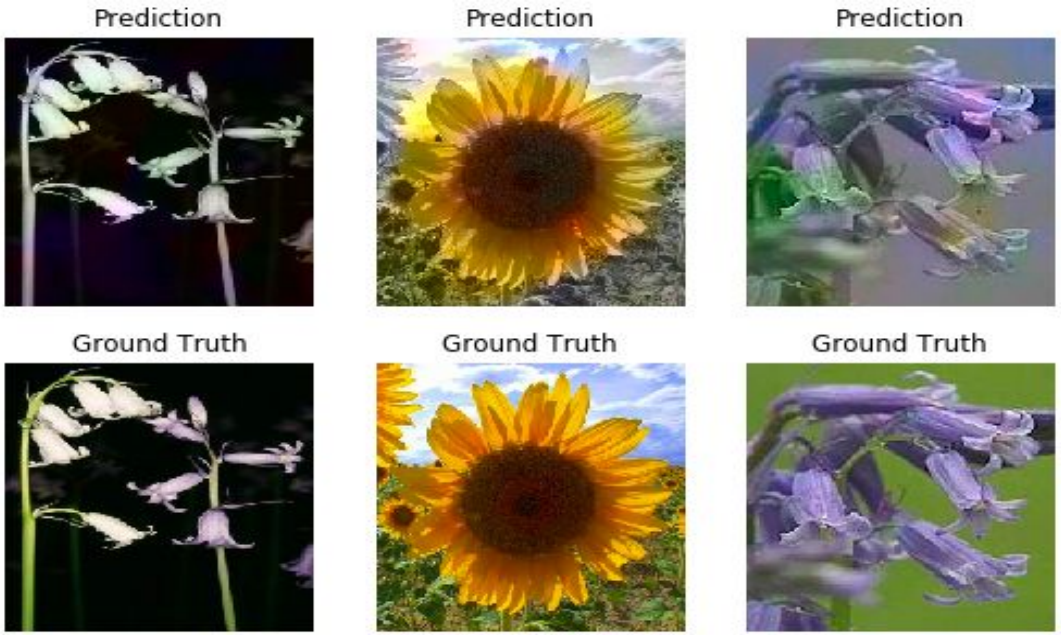
FOLIAGE

Results

MANMADE

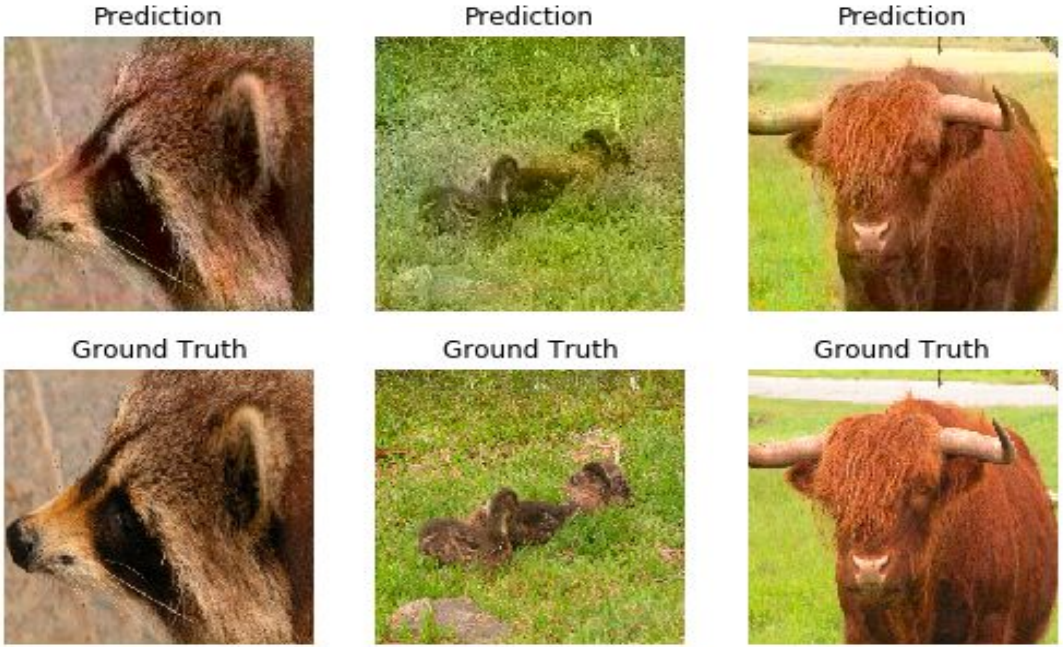
Results

FLOWER



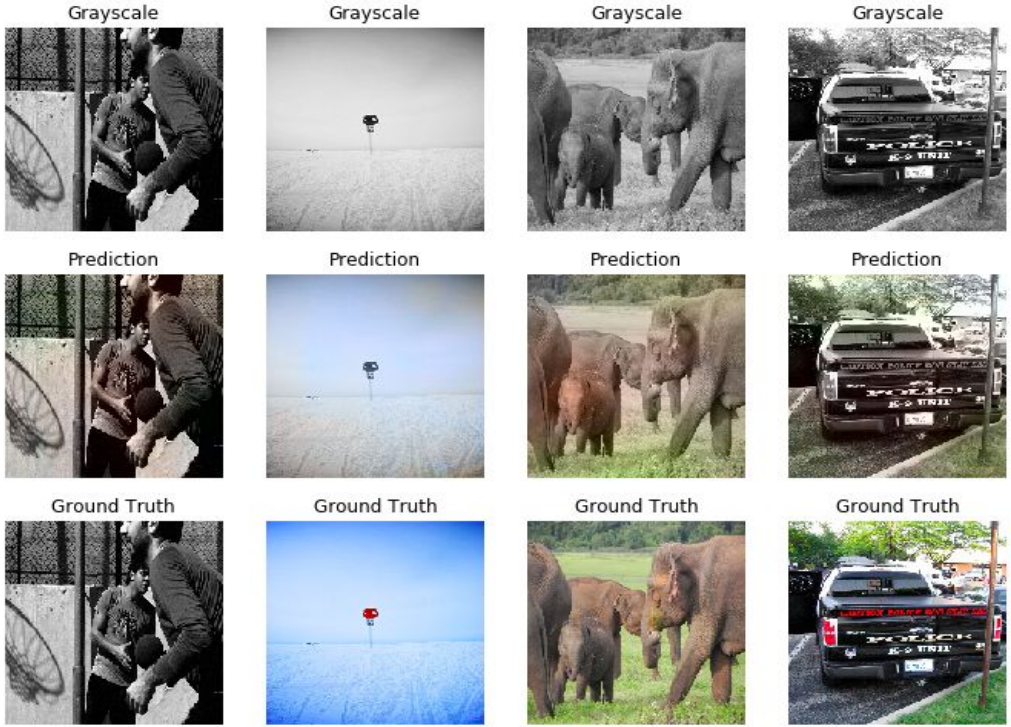
Results

ANIMAL



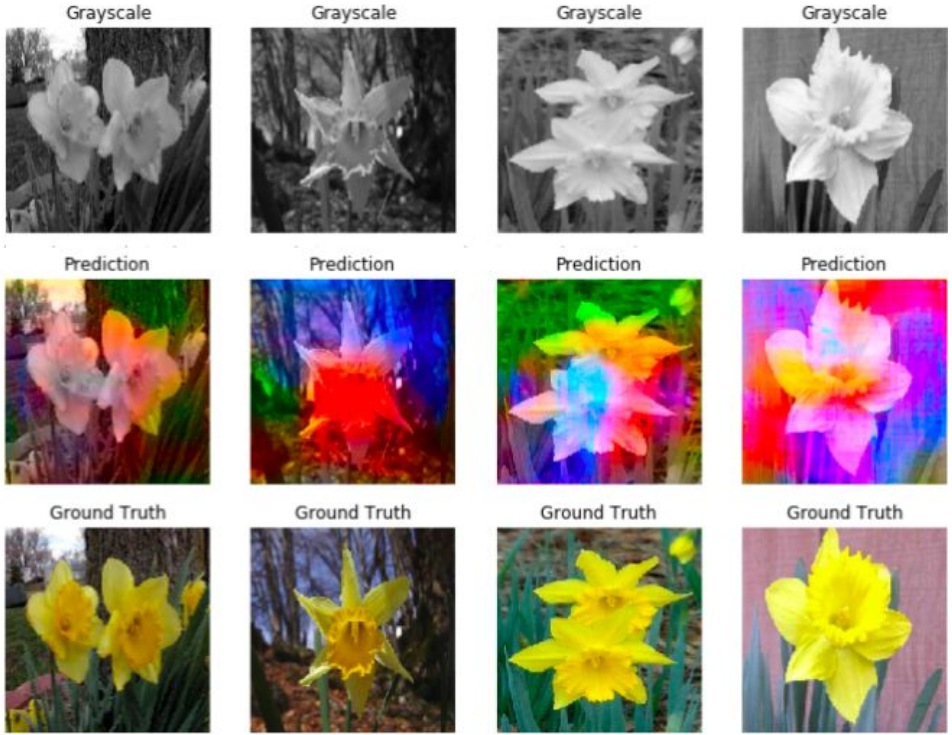
Results

DATASET: 10K **EPOCH: 100**



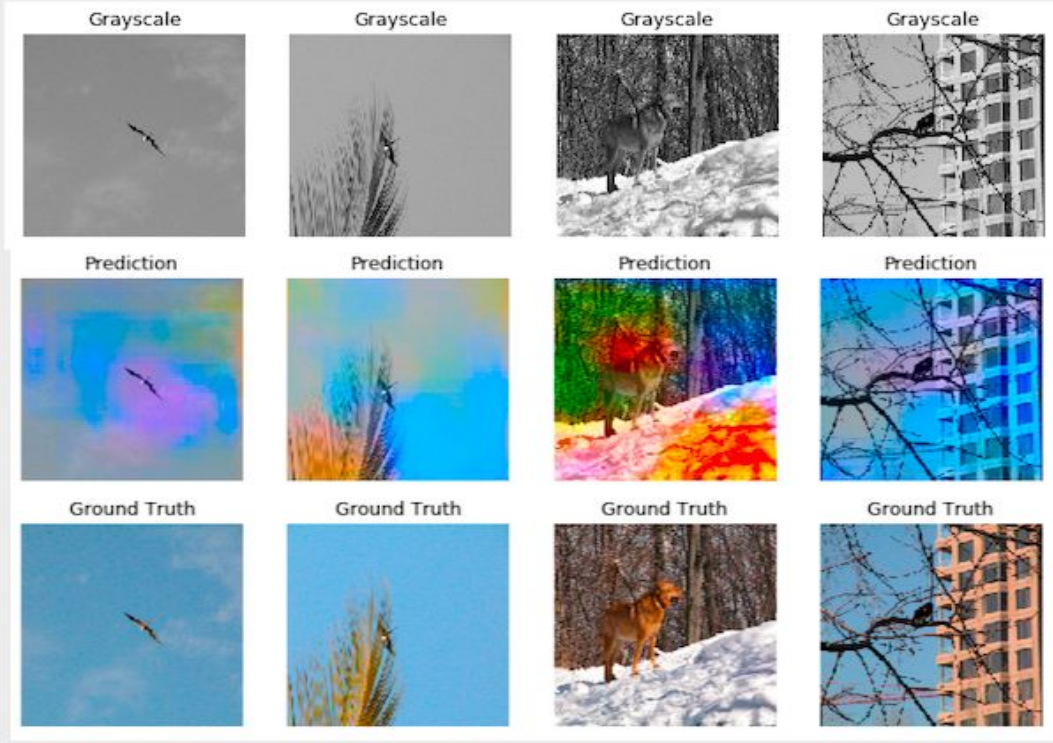
Results

DATASET: 10K **EPOCH: 500**

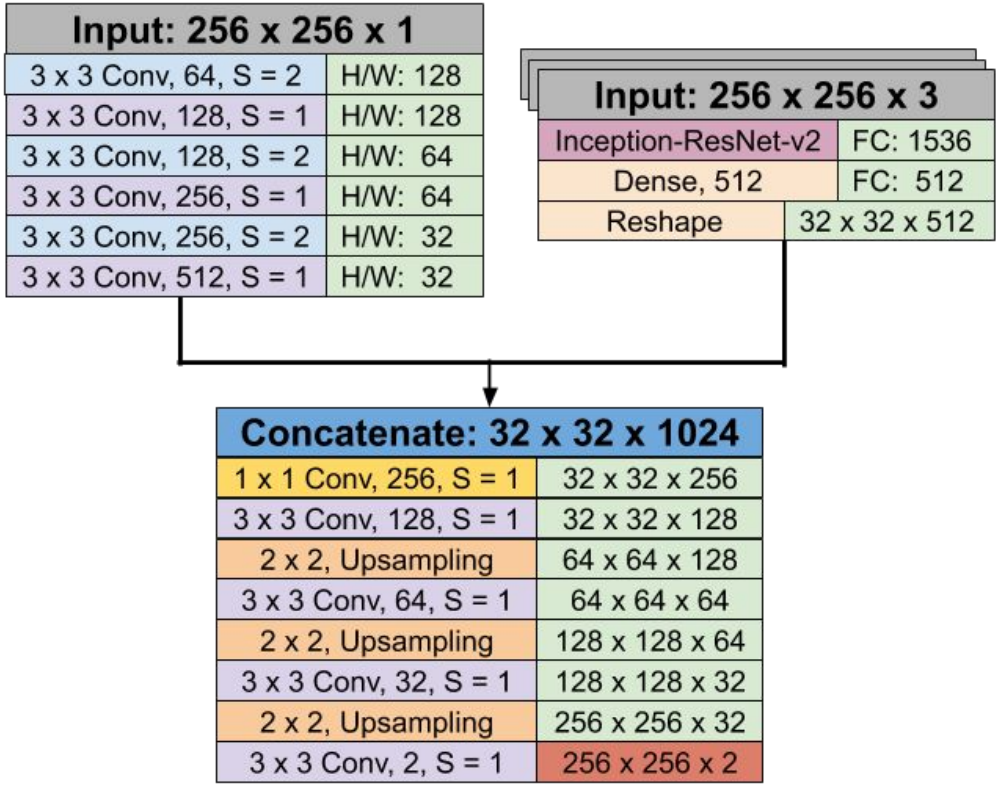


Results

DATASET: 16K **EPOCH: 200**



Transfer Learning: Inception-ResNet-V2



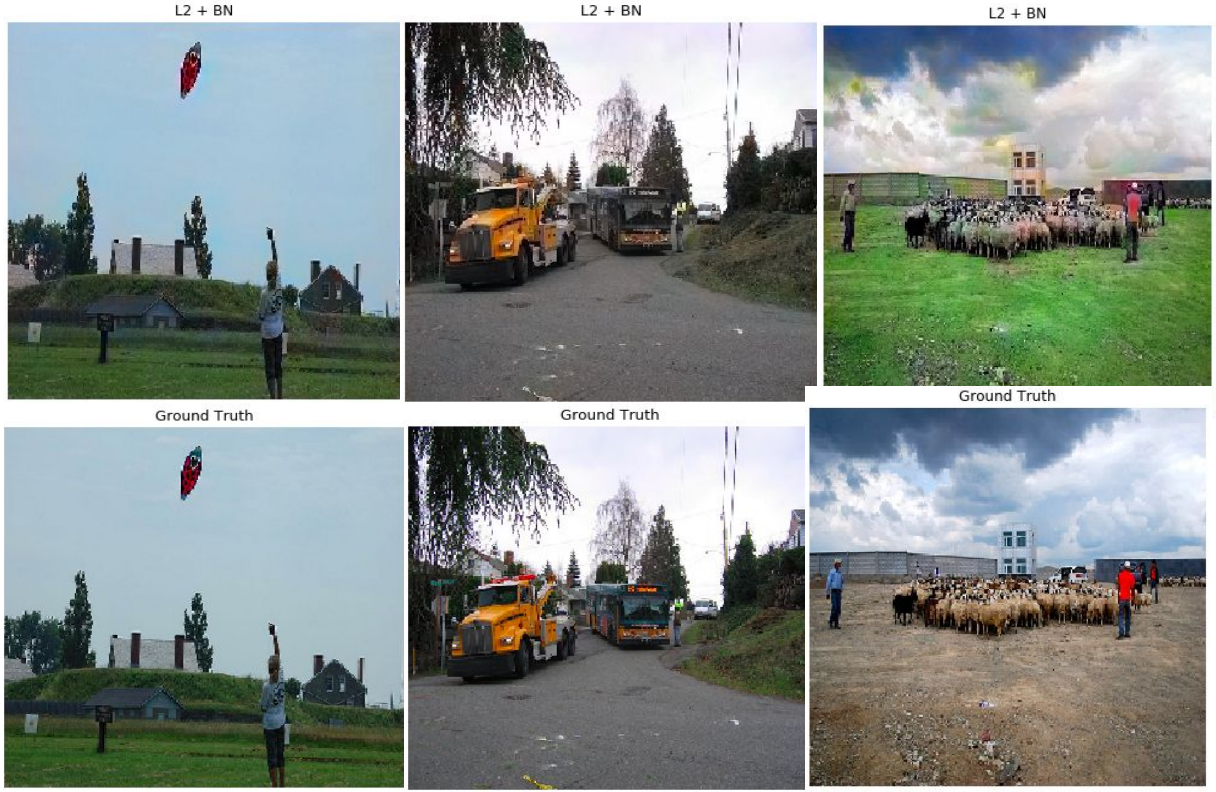
- **L2 Loss**
- **ReLU / TanH**
- **Batch Normalization**
- **No Dropout**
- **He Normal Weights**

Results - Nature (1 K)

Epochs = 100

	L2 LOSS	MAE
TRAIN	0.0166	0.0908
VAL	0.0859	0.1701
TEST	0.0179	0.0938

Results (Success) – VG (10 K)

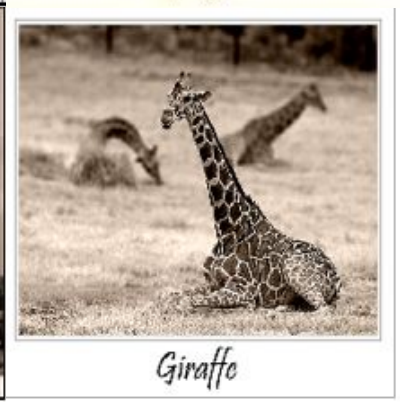


Prediction

Ground Truth

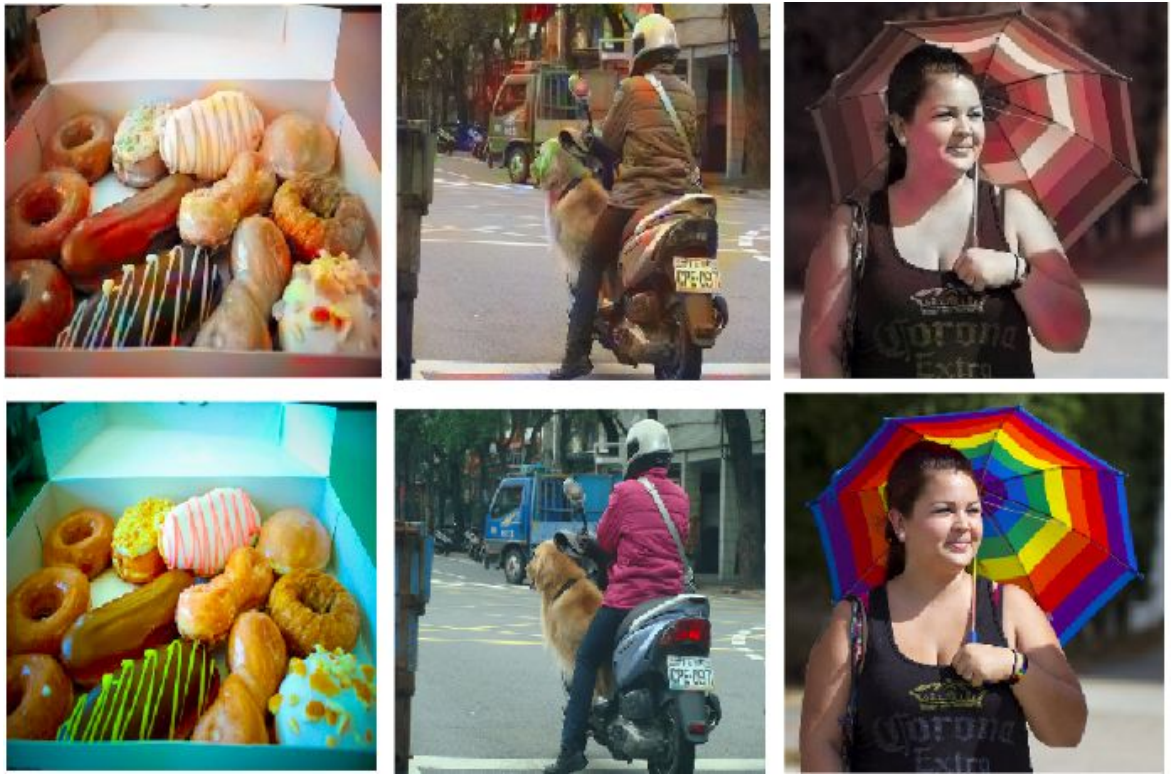
Results (Success) – VG (10 K)

Prediction



Ground Truth

Results (Failure) – VG (10 K)



Prediction

Ground Truth

Takeaways

- Size and type of Dataset important
- Regression - Sepia Tones for Multi-Color options
- Expansion - Classification and GAN models

Thank You!
Questions?