Colorful Image Colorization

Team 21:
Aakash Jain - 2019101028
Abhayram Nair - 2019102016
Aravind Narayanan - 2019102014
Ritvik Kalra -2019115002

Table of contents

01

Objective

You can enter a subtitle here if you need it

Objective

Produce plausible colorization

We want to create a system that can use semantic and texture cues to create plausible colorizations. In that sense, we want to model the dependency between semantic and texture cues with color.

Self-supervised learning technique

We aim to employ a self-supervised learning technique not requiring manual annotations or labels. We plan to use CNNs to predict a and b color channels from the lightness channel L (CIE).

Visually compelling results

We aim to generate visually compelling results that are able to fool human observers. Even if the images don't match the ground truth, if it is able to fool a human observer, we consider it a success.

Prior Work

Parametric

Non-parametric methods, given an input grayscale image, first define one or more color reference images to be used as source data.

Non-Parametric

Parametric methods learn prediction functions from large datasets of color images at training time, posing it as either a regression or classification problem.

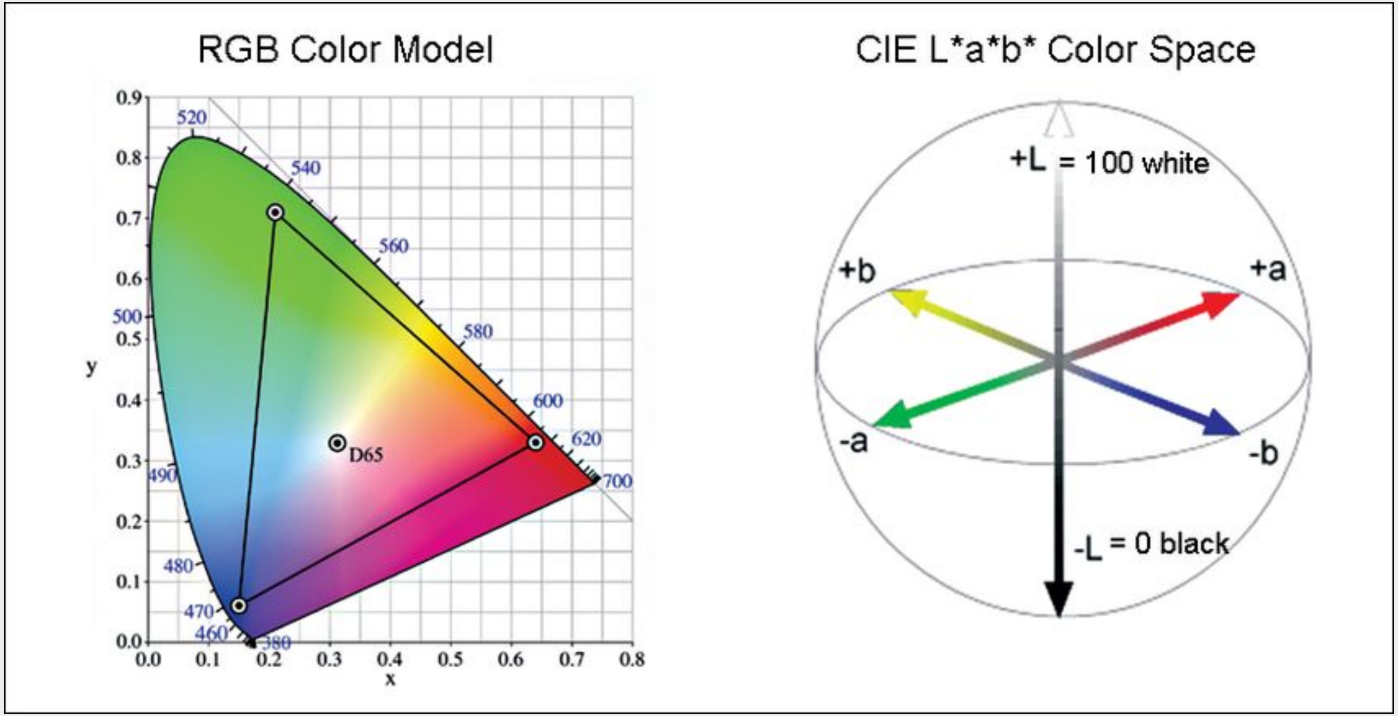
02

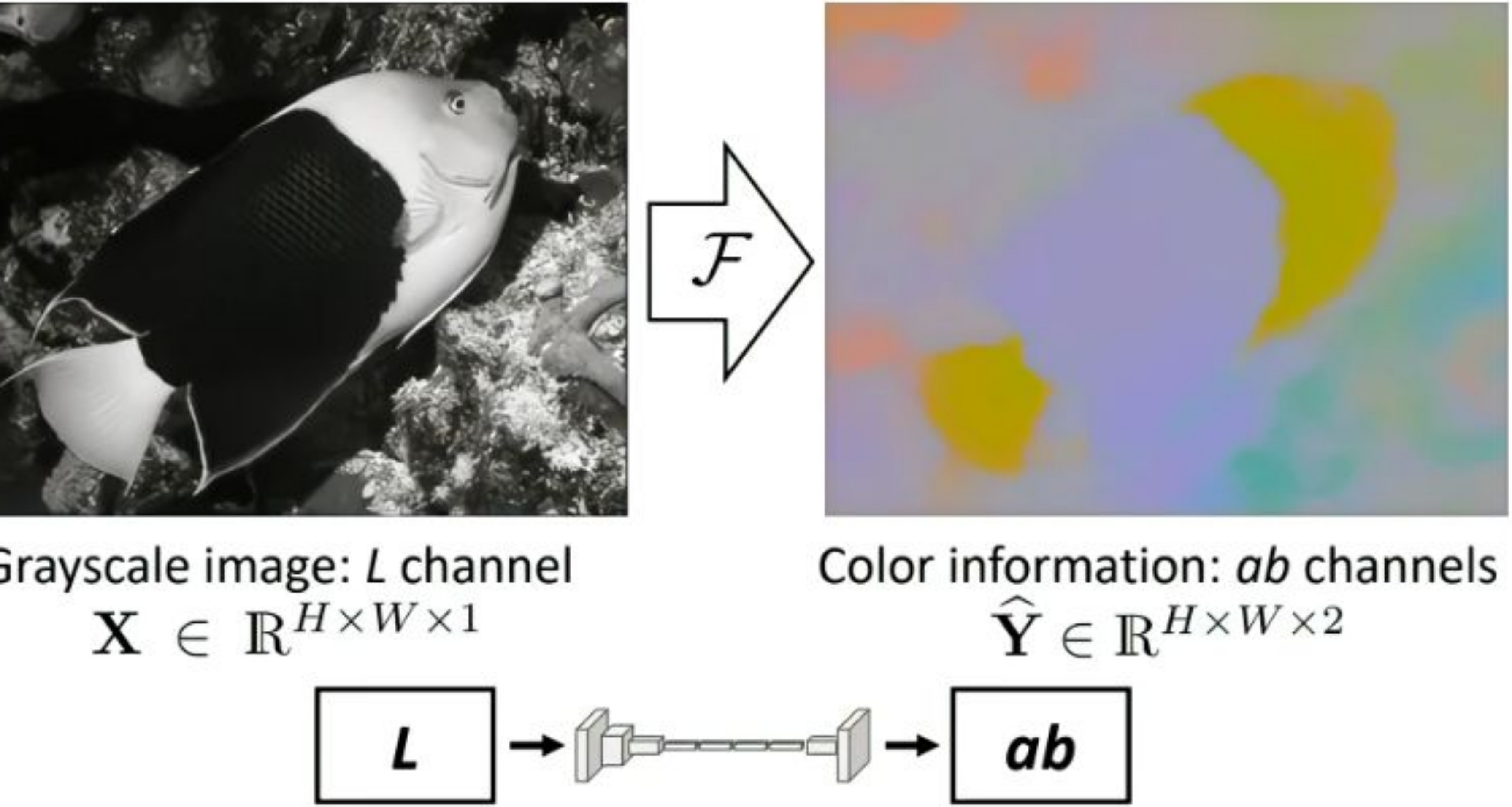
Methodology

Solution Approach of the paper

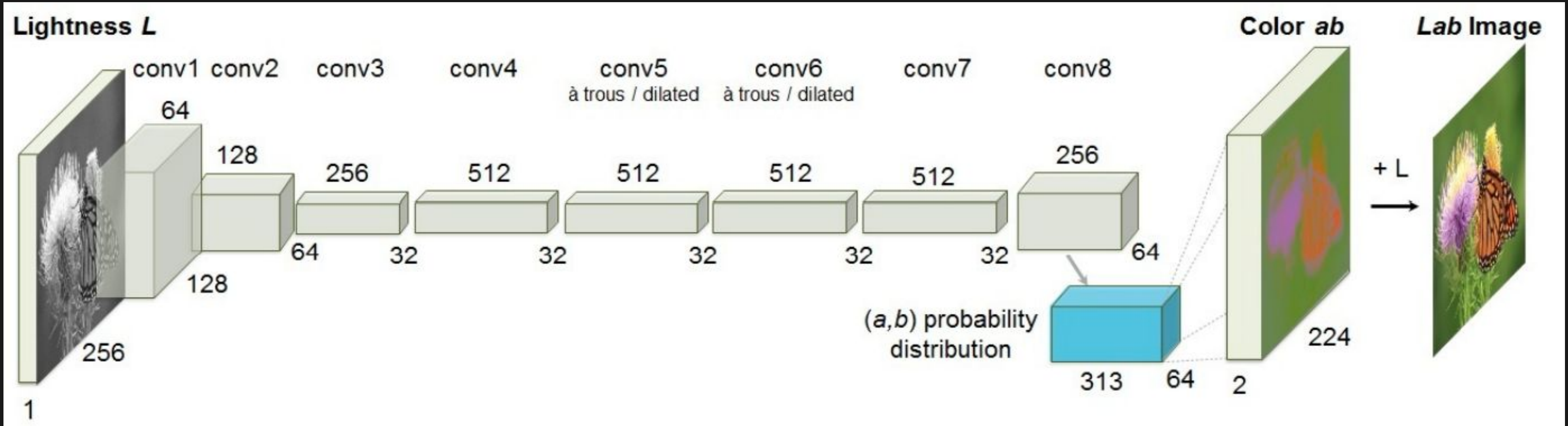
Preprocessing

Conversion from RGB Color Space to CIE Lab Space




$\mathcal{F}$
Grayscale image: *L* channel
$\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$
Color information: *ab* channels
$\widehat{\mathbf{Y}} \in \mathbb{R}^{H \times W \times 2}$
L
ab

Model Architecture



Loss Function

Multinomial Classification Loss

We arrive at this loss function after quantizing the ab output space into bins

$$L(\widehat{\mathbf{Z}}, \mathbf{Z}) = -\frac{1}{HW} \sum_{h,w} \sum_{q} \mathbf{Z}_{h,w,q} \log(\widehat{\mathbf{Z}}_{h,w,q})$$

Euclidean Loss

Not robust. Graying results. Implausible results if plausible colorizations are not convex.

$$L_2(\widehat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2} \sum_{h,w} \|\mathbf{Y}_{h,w} - \widehat{\mathbf{Y}}_{h,w}\|_2^2$$

Loss after class rebalancing

Removes the biasing seen towards desaturated colors through reweighting

$$L_{cl}(\widehat{\mathbf{Z}}, \mathbf{Z}) = -\sum_{h,w} v(\mathbf{Z}_{h,w}) \sum_{q} \mathbf{Z}_{h,w,q} \log(\widehat{\mathbf{Z}}_{h,w,q})$$

Class Probabilities to Point Estimates

Using Mode of predicted distribution

Vibrant but spatially inconsistent results

Using Mean of predicted distribution

Spatially consistent but desaturated results

$$\mathcal{H}(\mathbf{Z}_{h,w}) = \mathbb{E}\big[f_T(\mathbf{Z}_{h,w})\big], \quad f_T(\mathbf{z}) = \frac{\exp(\log(\mathbf{z})/T)}{\sum_q \exp(\log(\mathbf{z}_q)/T)}$$

Best of both worlds using Annealed mean operation

Interpolation is done by adjusting the temperature T of the softmax distribution and taking the mean of the result. Temperature T = 0.38 captures the vibrancy of mode while maintaining the spatial coherency of the mean

03

Tasks

Series of tests conducted as a part of the project

Tasks

Recreate Original Results

Implement the paper's solution approach from scratch on a smaller dataset and compare results and scores

Ablation Study

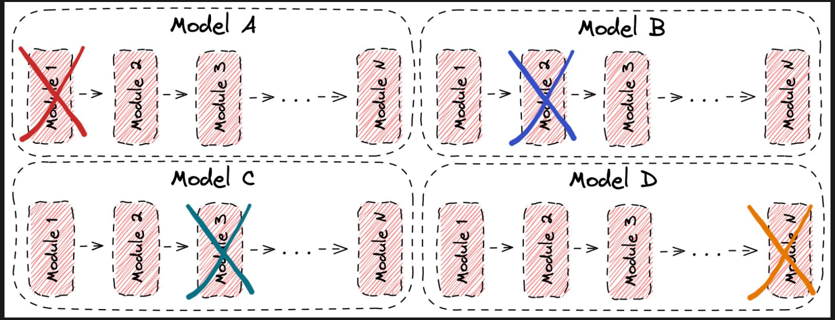
Analyse the importance of dilation convolution, different CNN layers by changing the number of layers and comparing results

Task Generalisation

Analyse the model's representative learning ability by conducting a study on task generalization.

Need for an Ablation Study

- Ablation study investigates the impact of changing the number of layers in a neural network.
- Systematically remove or add layers to determine the optimal number of layers that balance model complexity and performance.
- Ablation study helps understand how each layer contributes to the model's performance and improves the neural network architecture.
- This experimentation leads to more efficient and accurate machine learning models.



Ablation Study

Experiment - 1

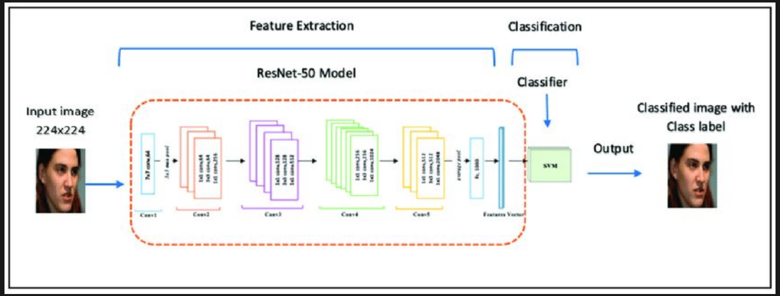
- 5 CNN Blocks
- 1 dilated convolutional block

Experiment - 2

- 4 CNN Blocks
- Executed without the dilated convolutional layers

Task Generalisation – Object Detection

- We aim to show that our colorized images are realistic enough to be useful for downstream tasks like object classification.
- To test our claim, we run a object classification pipeline using resnet50 with a pretrained model and parse our colourised images and compare the output.



04

Results

Results obtained from the tasks

Task 1: Implementation of Baseline Colorization

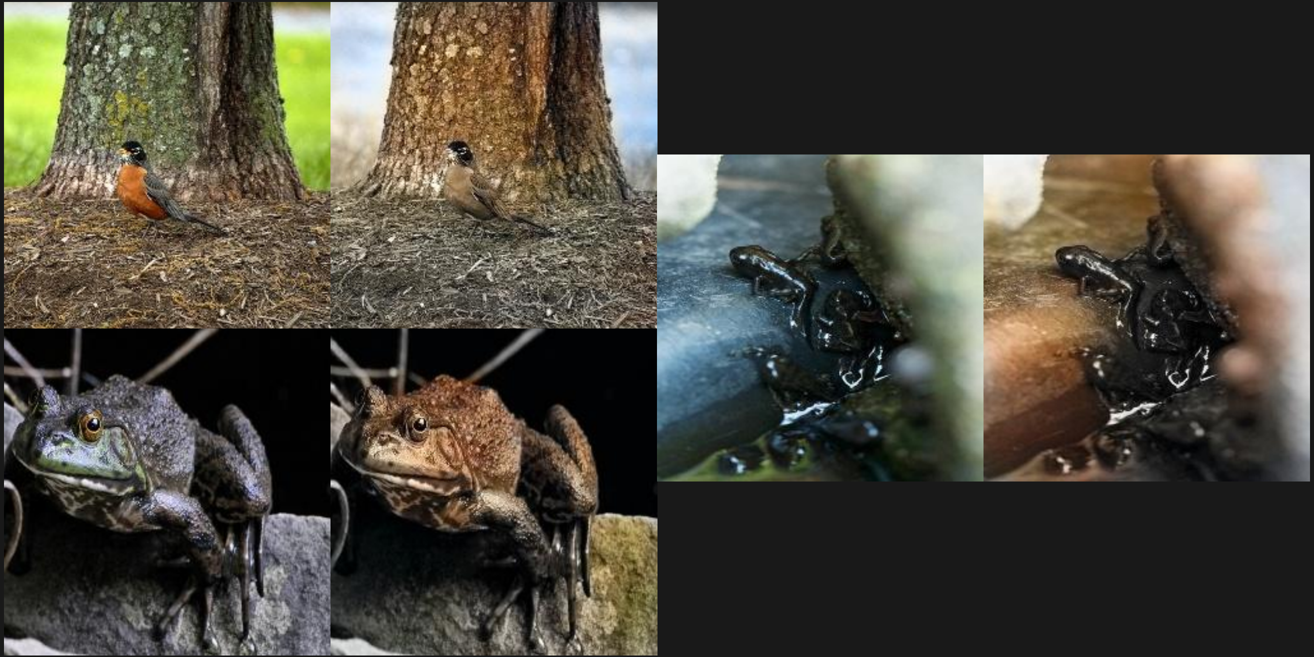
The results for the the colorization of images are categorised into the following categories:

- **Great looking results**
- **Bad looking results** (wrong colourization, faded/negligible colouring for obvious images)
- **Confusing results** (images whose colorisation is just as likely as ground truth and can confuse humans)
- **Color Bleeding results**
- **Legacy black** (images for which there is no colorized ground truth)

Good Looking Results

Badly Colorised Results

Results that can fool humans

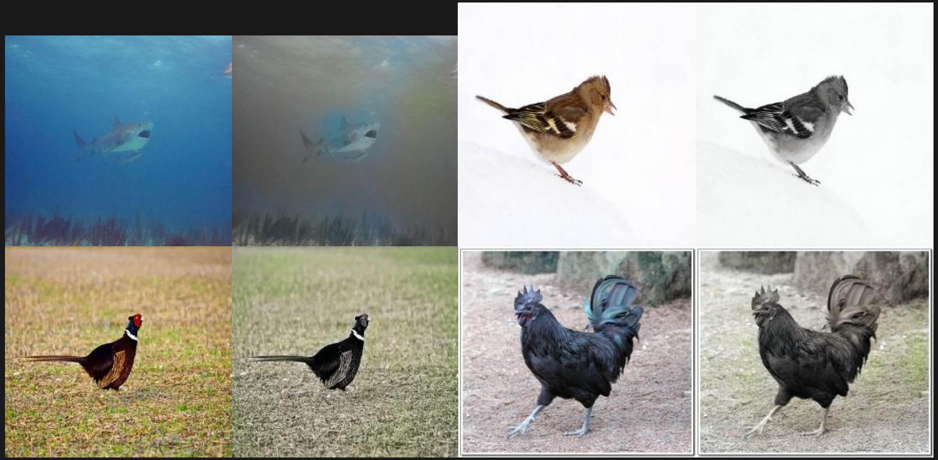


Color Bleeding Results

Legacy Black Image Results

Task 2: Ablation Study Experiment 1

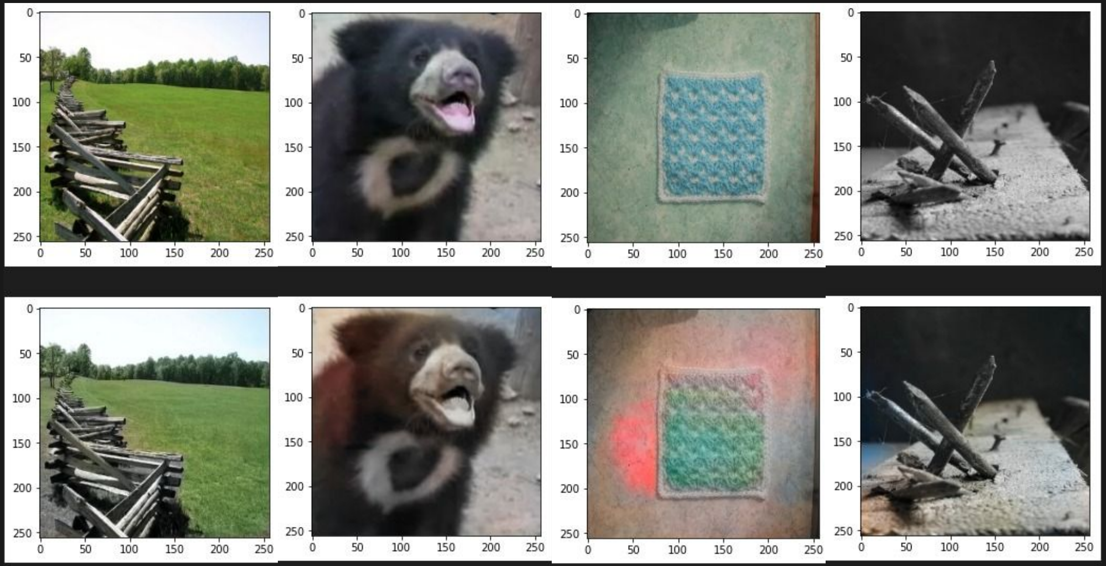
- A few sample ablation results obtained are as follows:



Task 2: Ablation Study
Experiment 2

- A few sample ablation results obtained are as follows:

Task 3: Image Classification



```
worm_fence 0.9999993
lumbermill 2.7462394e-07
picket_fence 2.1672166e-07
stone_wall 7.8555864e-08
cannon 5.1384337e-08
```

```
sloth_bear 0.8310369
lesser_panda 0.059745934
giant_panda 0.050954733
gibbon 0.019803464
American_black_bear 0.012891338
```

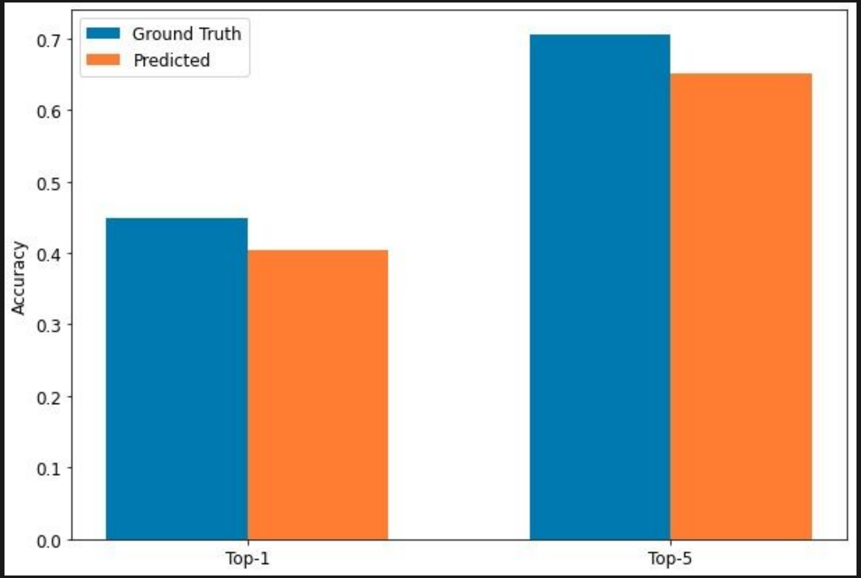
```
doormat 0.22281262
tray 0.17279541
milk_can 0.09955564
prayer_rug 0.07435007
washbasin 0.034598205
```

```
nail 0.89224887
sundial 0.017181283
horned_viper 0.015135847
wreck 0.005819217
pencil_sharpener 0.005610763
```

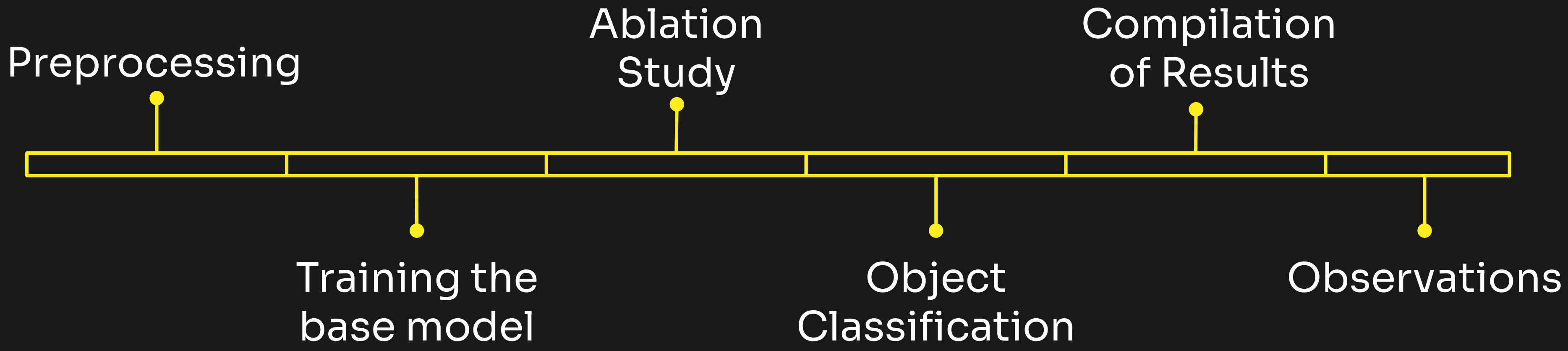
Performance

Model	Epoch	Test Loss	Validation Loss
Baseline Model	126	2.5449	3.7656
Ablation Task 1	50	3.2005	3.2157
Ablation Task 2	50	3.1573	3.1835

Performance Metrics



Timeline


Preprocessing
Training the base model
Ablation Study
Object Classification
Compilation of Results
Observations

Thanks!