Image Colorization

COMPUTER VISION PROJECT

Submitted By

Harsh Parashar	Prateek Singhal	Amit Panwar
M22AIE210	M22AIE215	M22AIE202

Project Objective

- Create a model for transforming grayscale pictures into vibrant and realistic colored pictures
- Treat colorization as a classification task and utilize class-rebalancing during training.
- Due to resource constraints, train the system on a minimal dataset of ~14k color images to achieve comparable color prediction quality.
- Develop an end-to-end API and web application to use the model as a SaaS offering.

Dataset

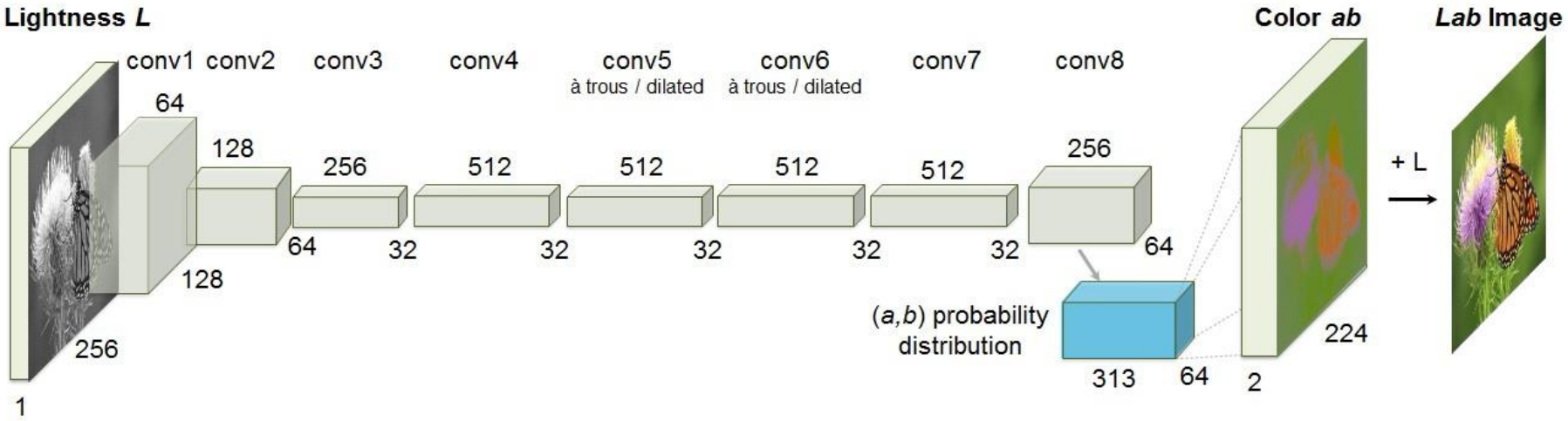
- We will be using combination of following dataset to include diversity in images and train a generalize image colorization model
  - Sport Image
    - Dataset contains ~14.4k sports images sized 224×224 covering 100 different sports.
  - Landscape Images
    - The dataset includes ~7.1k landscape images 150×150 of trees, buildings, mountains, glaciers, and trees.
  - Pascal VOC 2012
    - Dataset contains ~31k images however we will only be using images categorized in Train set as our corpse i.e. ~17k images. The dataset include images for various sizes varying from 250×500 to 500×500 (this is approximate size range got from random check)
- We are taking 30% from each above dataset and create our working dataset with ~11k Samples

Data Preprocessing

- Convert the colored images to grayscale
- Normalize grayscale images and colored images between 0 and 1 for easier processing.
- Resize images to 256 X256.
- Shuffle the dataset to remove any inherent order.
- Split dataset into training, validation, and test sets (70%, 10% and 20% respectively).




Model Architecture



Convergence Approach

- **L2 Loss functi**on: It is Euclidian distance between predicted and the ground truth (but this is not robust as it may cause desaturated results due to averaging)
- **Class Rebalance**: It helps in rebalancing the ab (color gamut) values which helps in getting vibrant images
- **Class Probabilities to Point Probabilities**: It helps in getting better accuracies in the color of objects

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

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

$$v(\mathbf{Z}_{h,w}) = \mathbf{w}_{q^*}, \text{ where } q^* = \arg\max_q \mathbf{Z}_{h,w,q}$$

$$\mathbf{w} \propto \left((1-\lambda)\widetilde{\mathbf{p}} + \frac{\lambda}{Q}\right)^{-1}, \quad \mathbb{E}[\mathbf{w}] = \sum_q \widetilde{\mathbf{p}}_q \mathbf{w}_q = 1$$

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

Evaluation Metrics

Evaluation Metrics used in paper	Evaluation Metrics used in our project
Raw accuracy (AuC)	Raw accuracy (AuC)
Semantic interpretability (VGG classication)	Semantic interpretability (VGG classication)
Perceptual realism (AMT)	

Semantic interpretability (VGG classification): This is achieved by feeding our fake colorized images to a VGG network that was trained to predict ImageNet classes from real color photos.

Raw accuracy (AuC): It is computed the percentage of predicted pixel colors within a threshold L2 distance of the ground truth in ab color space

Perceptual realism (AMT): This is a real vs. fake two-alternative forced choice experiment on Amazon Mechanical Turk (AMT).

Colorization Result on ImageNet

Colorization Results on ImageNet							
Method	Model Params (MB)	Model Feats (MB)	Runtime (ms)	AuC non-rebal (%)	AuC rebal (%)	VGG Top-1 Class Acc (%)	AMT Labeled Real (%)
Ground Truth	–	–	–	100	100	68.3	50
Gray	–	–	–	89.1	58.0	52.7	–
Random	–	–	–	84.2	57.3	41.0	13.0±4.4
Dahl [2]	–	–	–	90.4	58.9	48.7	18.3±2.8
Larsson et al. [23]	588	495	122.1	**91.7**	65.9	**59.4**	**27.2±2.7**
Ours (L2)	129	127	17.8	91.2	64.4	54.9	21.2±2.5
Ours (L2, ft)	129	127	17.8	91.5	66.2	56.5	23.9±2.8
Ours (class)	129	142	22.1	91.6	65.1	56.6	25.2±2.7
Ours (full)	129	142	22.1	89.5	**67.3**	56.0	**32.3±2.2**

Project Plan

S.No	Tasks	Owner	Status	Remarks
	Required Plan			
1	Repository Creation	Harsh Parashar	**Complete**	
2	Data Gathering	Harsh Parashar	**In-Progress**	
3	Data Preprocessing code	Harsh Parashar	**In-Progress**	
4	Architecture Coding	Prateek Singhal	**In-Progress**	
5	Evaluation Metric Coding	Amit Panwar	In-Progress	
6	API Creation	Amit Panwar	Not-Started	
7	Web Application Creation	Harsh Parashar	Not-Started	
	Aspirational Plan			
8	Model	Harsh/Amit/Prateek	Not-Started	

Thank You!