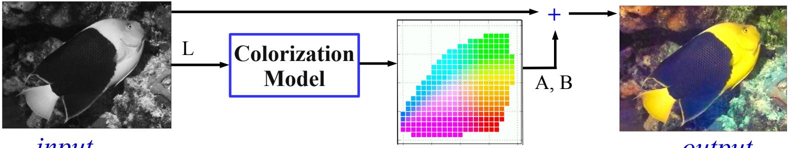
Automatic Image Colorization using Generative Adversarial Networks

Team: Yet Another Layer - YAL

Cameron Fabbri, Md Jahidul Islam

Image Colorization

Problem: produce a *realistic* coloring of gray-scale images



input *generated color channels* *output*

Applications

Colorizing old photos and movies

Colorizing underwater images

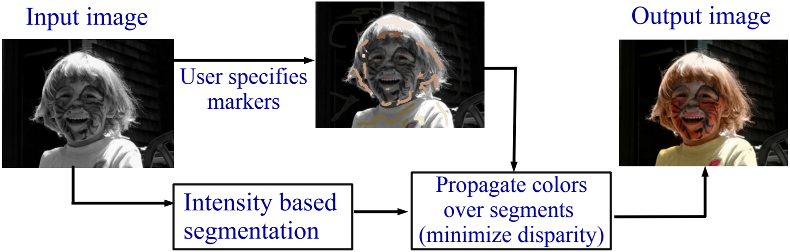
Figure sources: [1], [2], [3]

Background

- Algorithmic choices:
  - **Image-to-image translation**
  - Classification
- Colorspace choices:
  - **LAB**, RGB
- Approaches
  - Classical
  - Deep learning based
    - Generative models
    - **Adversarial model**

Classical Approaches

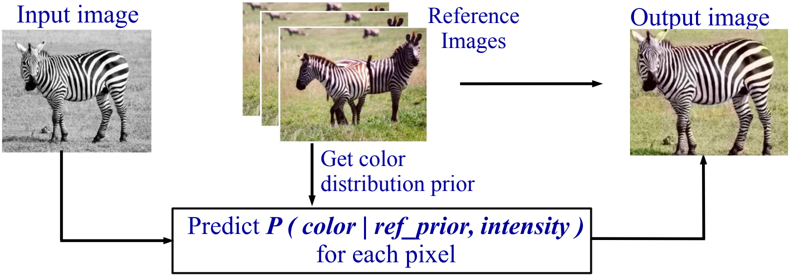
Colorization using User-specified Prior [4]



Pros	**Cons**
Fast	Not automatic
Reliable	Not scalable

Classical Approaches

Colorization using Multi-modal Prediction [5]



Pros

Automatic

Handles multimodality

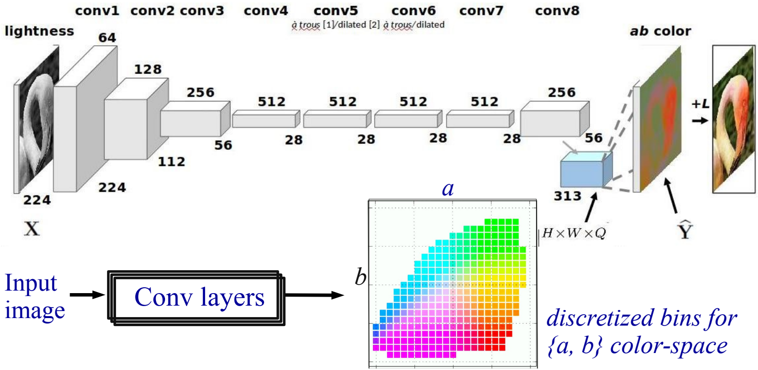
Cons

Depends heavily on reference prior

Not practical

Generative Approach

Colorful Image Colorization model [1]



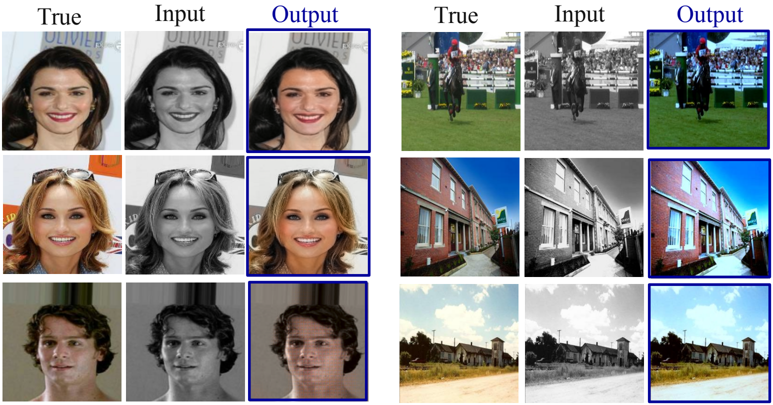
Colorful Image Colorization model [1]

Adopted Changes

Model Changes	Original model	Adopted model
Output	Classification Probabilities	Image
Loss function	Cross-entropy based $L_{col}(Z, \hat{Z}) = -\sum_p Z_p \sum_{q \in Q} Z_p[q] log(\hat{Z}_p[q])$	L1, L2 $L_1 = \lambda \sum_p \lvert I_{AB}[p] - \hat{I}_{AB}[p] \rvert$ $L_2 = \lambda \sum_p (I_{AB}[p] - \hat{I}_{AB}[p])^2$
Trained on	ImageNet	CelebA, Places2
Implementation	Caffe	Tensor-flow
Training	Not end-to-end	End-to-end

Results

Colorful Image Colorization model as generator only



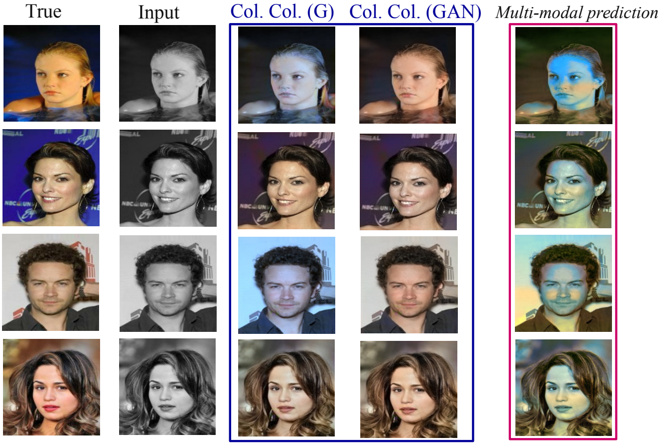
CelebA dataset

Places2 dataset

Loss functions: **L1 loss**, L2 loss, Least-squared loss

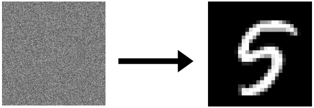
Results

Colorful Image Colorization model as generator in a GAN



Generative Adversarial Networks (GANs)

- Two player minimax game
  - Discriminator **D**
  - Generator **G**
- D is trained to discriminate between a real image and a generated image



- G is trained to generate an image that will fool D
- Both G and D are neural networks

$$\min_{G} \max_{D} \mathbb{E}_{x \sim p_{data(x)}}[log D(\mathbf{x})] + \mathbb{E}_{z \sim p_z(z)}[log(1 - D(G(z)))]$$

- Conditional GANs generate images given conditional information, such as a class label.

GAN Variations

- Deep Convolutional GANs (DCGANs)
  - Bridge the gap between GANs and Deep Learning
- Least Squares GANs (LSGANs)
  - Use a least squares loss for the discriminator
- Energy-Based GANs (EBGANs)
  - Model the discriminator as an energy function

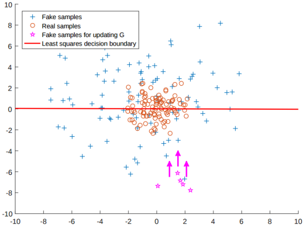


Fig. source: [7]

- Wasserstein GAN (WGAN)
  - Minimizes the Earth Mover distance between two distributions

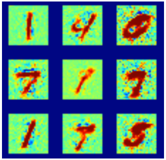
GAN Comparison Results

<u>MNIST</u>
GAN, DCGAN

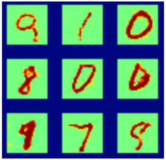
<u>CelebA</u>
EBGAN, WGAN

<u>LSUN</u>
LSGAN

GAN



DCGAN



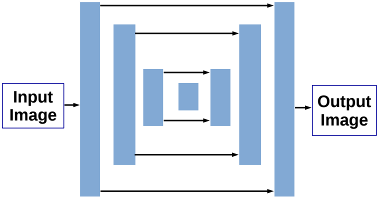
EBGAN

LSGAN

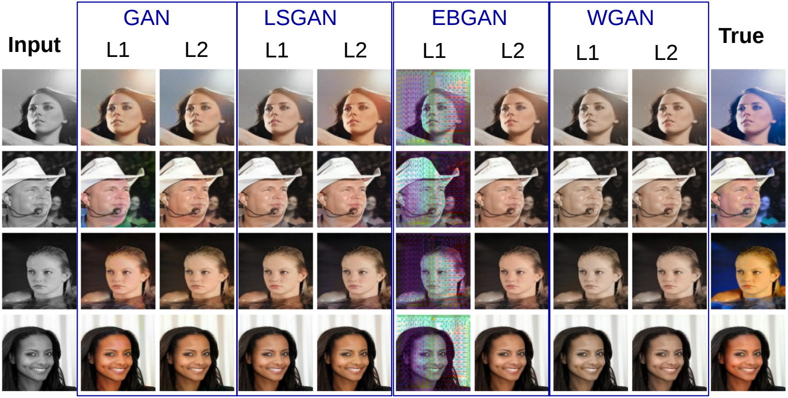
WGAN

GANs Towards Colorization

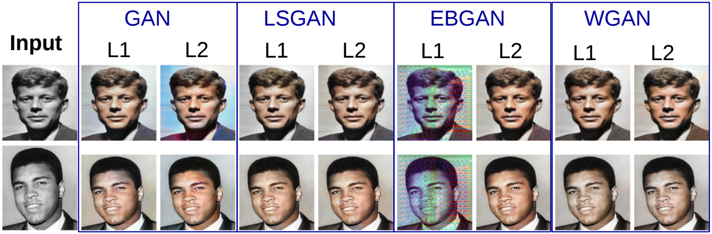
- Tested two generators
  - Pix2Pix
  - Colorful Image Colorization
- Pix2Pix generator is modeled as an encoder-decoder with skip connections
- Combine L1 and L2 loss with GAN Variatons



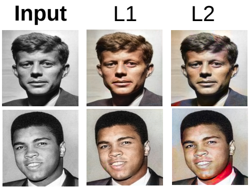
Results in Comparison



Legacy Grayscale Results



Pic2Pix model as Generator



Colorful Image Colorization model as Generator

Conclusions and Lessons Learned

- Automatic colorization is difficult due to
  - Ambiguity
  - Multi-modality
- Need large datasets to *learn* colorization
- GANs are tricky
- Long turnover time, at least 1 to 2 days
- Our model
  - End to end trainable
  - Inference is 30fps on CPU, 60fps on GPU

References

1. Richard Zhang, Phillip Isola, and Alexei A Efros. *Colorful image colorization*. In European Conference on Computer Vision, pages 649666. Springer, 2016.
2. Huimin Lu, Yujie Li, and Seiichi Serikawa. *Underwater image enhancement using guided trigonometric bilateral filter and fast automatic color correction*. In Image Processing (ICIP), 2013 20th IEEE International Conference on, pages 34123416. IEEE, 2013.
3. Luz A Torres-Mendez and Gregory Dudek. *Color correction of underwater images for aquatic robot inspection*. In International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, pages 6073. Springer, 2005.
4. Anat Levin, Dani Lischinski, and Yair Weiss. *Colorization using optimization*. In ACM Transactions on Graphics (ToG), volume 23, pages 689694. ACM, 2004.
5. Guillaume Charpiat, Matthias Hofmann, and Bernhard Scholkopf. *Automatic image colorization via multimodal predictions*. Computer VisionECCV 2008, pages 126139, 2008.
6. Goodfellow, Ian, et al. *Generative adversarial nets*. Advances in neural information processing systems. 2014.
7. Mao, Xudong, et al. *Least squares generative adversarial networks*. arXiv preprint ArXiv:1611.04076 (2016).
8. Arjovsky, Martin, Soumith Chintala, and Lon Bottou. *Wasserstein gan*. arXiv preprint arXiv:1701.07875 (2017).
9. Zhao, Junbo, Michael Mathieu, and Yann LeCun. *Energy-based generative adversarial network*. arXiv preprint arXiv:1609.03126 (2016).
10. Radford, Alec, Luke Metz, and Soumith Chintala. *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks*. ICLR (2016): https://arxiv.org/abs/1511.06434.