Image Colorization with Convolutional Neural Networks

Div Dasani, Northwestern University
Advised by Ying Wu, Professor of Electrical and Computer Engineering, Northwestern University

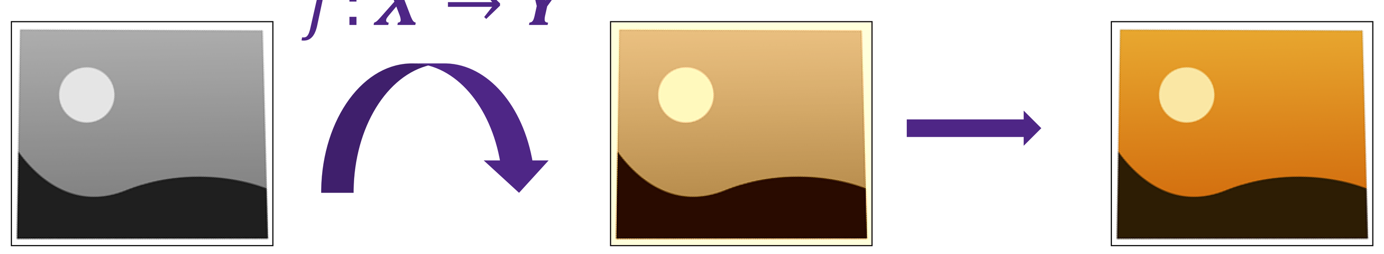
Northwestern University

I. Goals

$$P = \begin{bmatrix} p_{1,1} & \cdots & p_{1,w} \\ \vdots & \ddots & \vdots \\ p_{l,1} & \cdots & p_{l,w} \end{bmatrix} \quad p_{i,j} = (L_{i,j}, a_{i,j}, b_{i,j})$$

$c_{i,j} = (a_{i,j}, b_{i,j})$,
$p_{i,j} = (L_{i,j}, c_{i,j})$

$f: X \to Y$

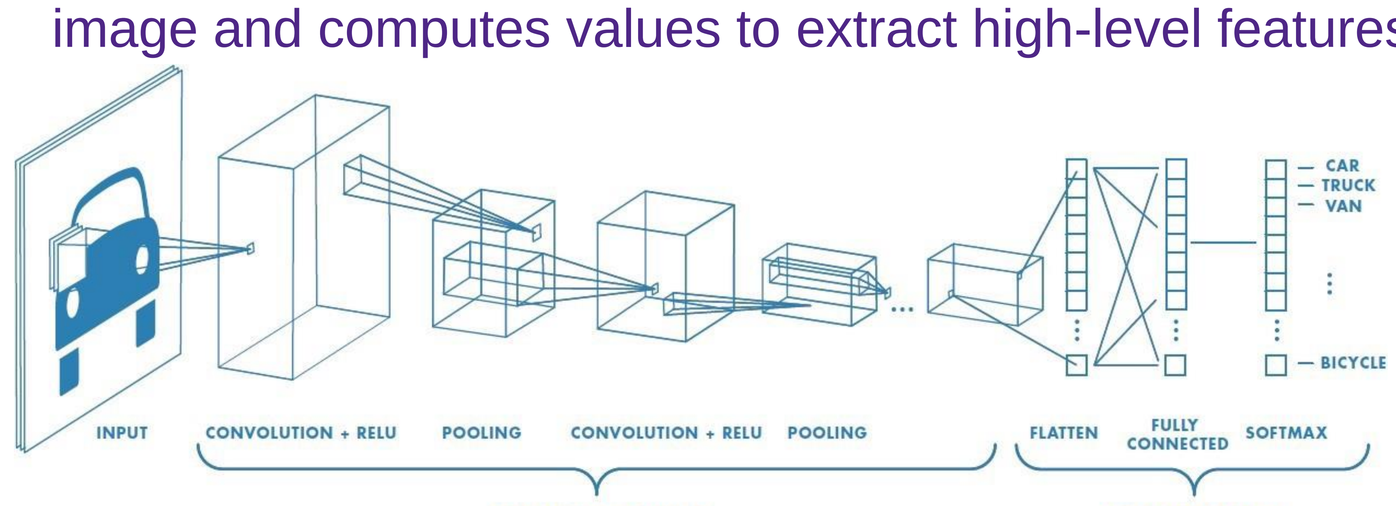


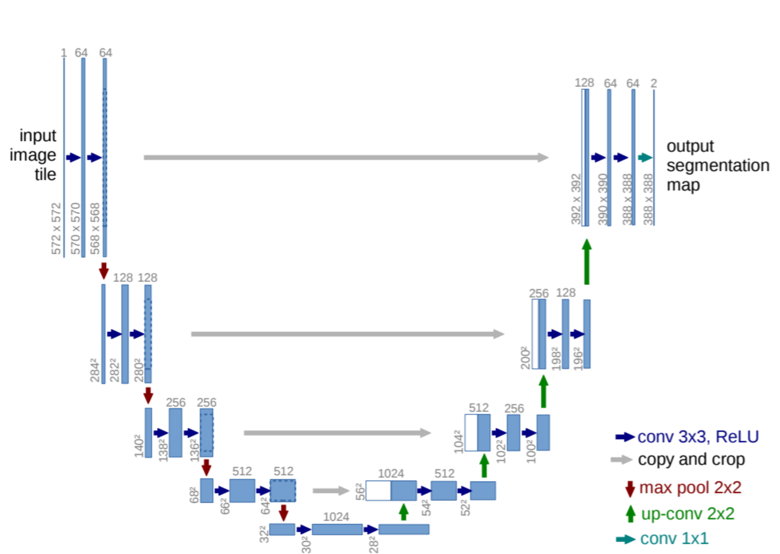
$$X \in \mathbb{R}^{l*w} \quad + \quad Y \in \mathbb{R}^{2*l*w} \quad = \quad [X;Y] \approx P$$

Image colorization:
- The process of recoloring (B&W) images authentically
- Images are made of pixels, which are represented in Lab space
- Goal: to consistently colorize images realistically with no context
  - Mathematically, retrieve $\widehat{p_{i,j}} \approx p_{i,j}\ \forall\, i,j\ \in l, w$ via some f

II. Model Architecture

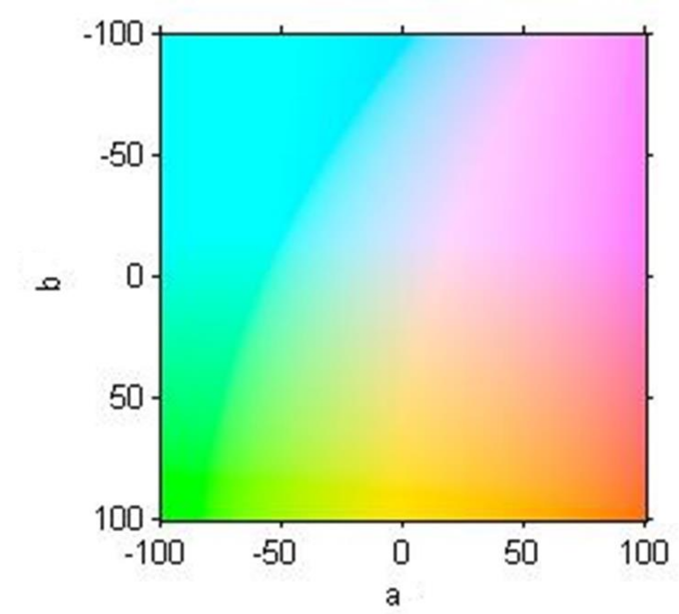
Convolutional Neural Networks:
- Set of algorithms, modeled after the brain, to recognize patterns
- Takes in an image and employs filters to learn characteristics
  - Initial filters primitive (diagonal lines)
  - deep filters are advanced (tire, headlight)
- This is done with a convolutional kernel, which moves along the image and computes values to extract high-level features





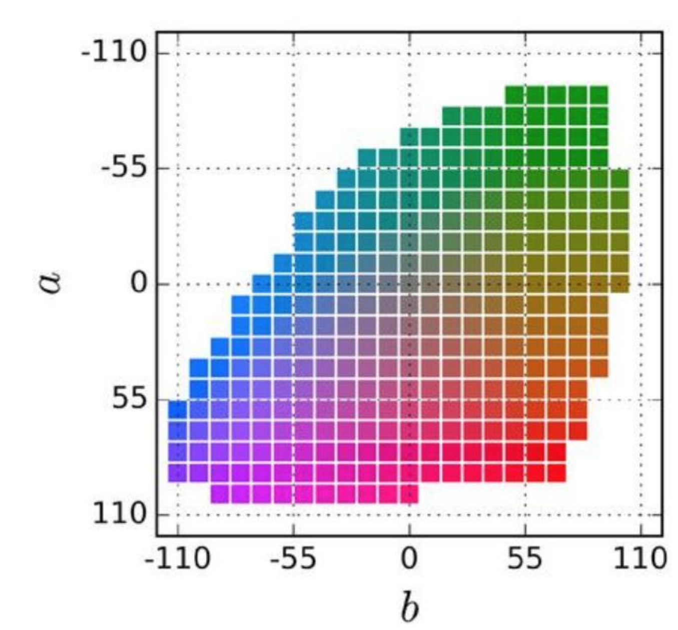
U-Net:
- Input progressively downsampled via contracting encoders and then upsampled with expansive decoders
- Feature learning in middle layers is compact
- Pooling layers are not employed

III. Regressive and Classifying Approaches



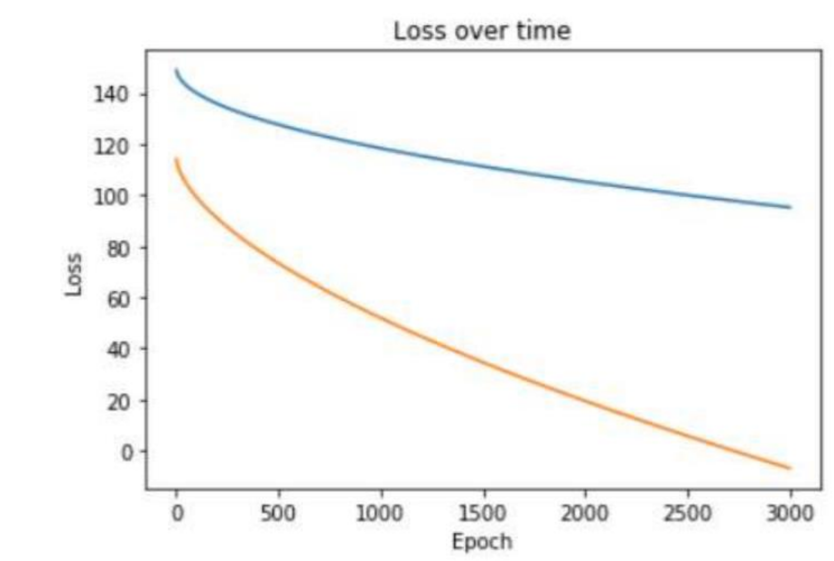
Regressive approach:
- Goal: predict $\widehat{c_{i,j}}$ by minimizing absolute difference between $\widehat{c_{i,j}}$ and $c_{i,j}\,\forall\, i,j\ \in l, w$
- Akin to minimizing the distance between two vectors in $\mathbb{R}^2$
- The loss function is given by,

$$L_R(X) = \frac{1}{n}\sum_{i=1}^{n} \left\| f(X)^i - Y^i \right\|_2^2$$



Classifying approach:
- The continuous color space is quantized into discrete bins, allowing each bin to be treated as a class.
- Task is now to select the optimal class for each pixel rather than distance minimization problem
- The loss function (with Q classes) is given by,

$$L_C(X) = -\frac{1}{n}\sum_{i=1}^{n}\sum_{q=1}^{Q} Y^{i,q} \log(f(X)^{i,q})$$



Experimentation:
The same CNN was trained using both the regressive and classifying approaches on the same batch of images. The loss over time for each approach is depicted above.

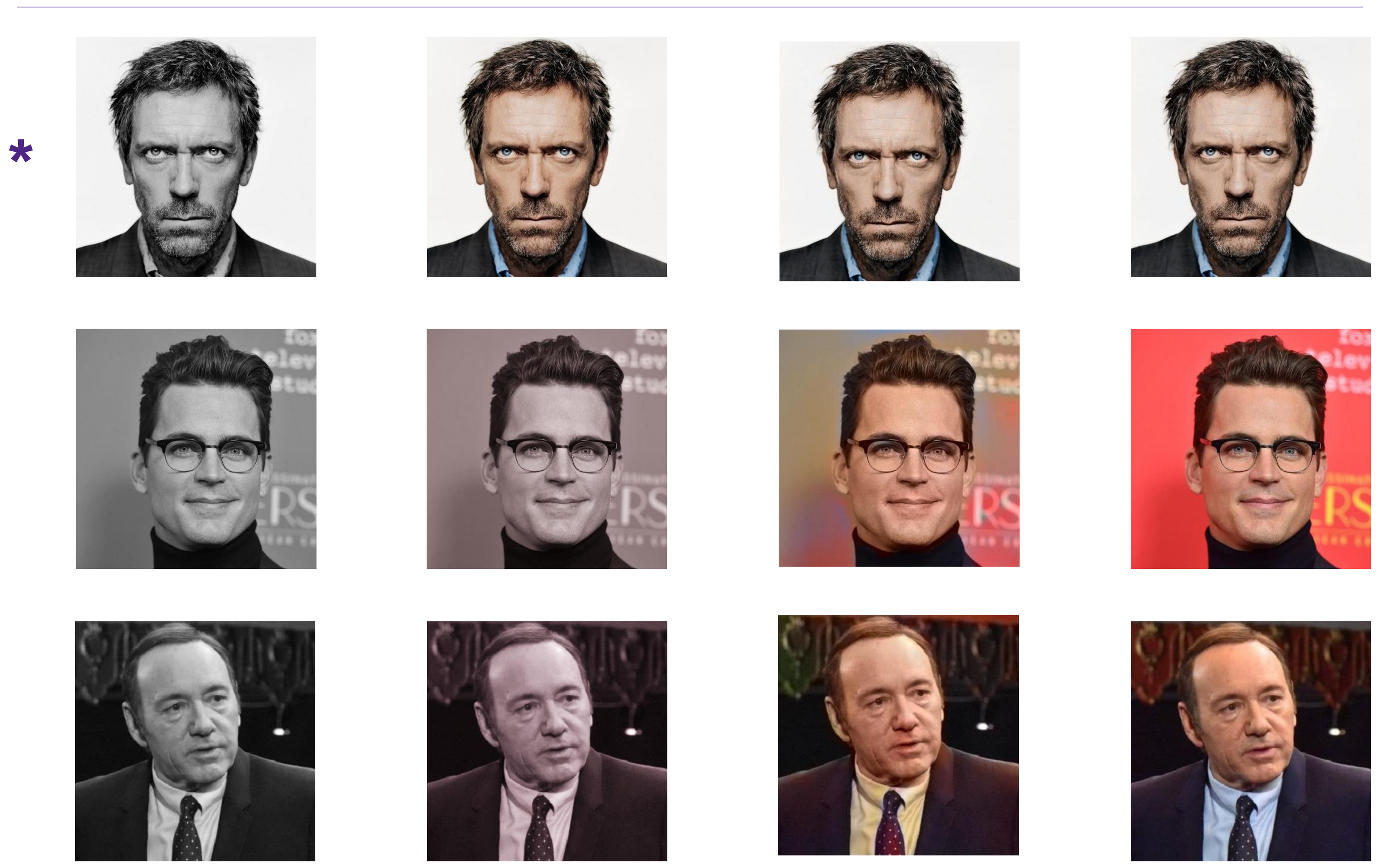
The stagnation of the former approach at a high loss is due to the "brown well" problem.

IV. Results

(* denotes training row)

Black & White	Regressive Model	Classifying Model	Ground Truth

*



V. Conclusion

- Image colorization is an underconstrained problem
- CNN's are powerful tools to solve unstructured problems
- Regressive approaches are risk-averse, thereby suffering from the "brown well" problem
- Coupling a classifying approach with a discretized color space renders vibrantly colored, visually appealing images

Next Steps:
- Creating a generalized colorization algorithm by training on a variety of images
- Making the algorithm more robust to different image sizes or media types (videos, gifs, etc.)

Thank you to Professor Wu for advising my research

- Ronneberger, et. Al: U-net: Convolutional networks for biomedical image segmentation (2015)
- Federico Baldassarre et. Al: Deep Koalarization: Image Colorization using CNNs and InceptionResnet-v2 (2017)
- Zhang et. Al: Real-Time User-Guided Image Colorization with Learned Deep Priors (2017)