

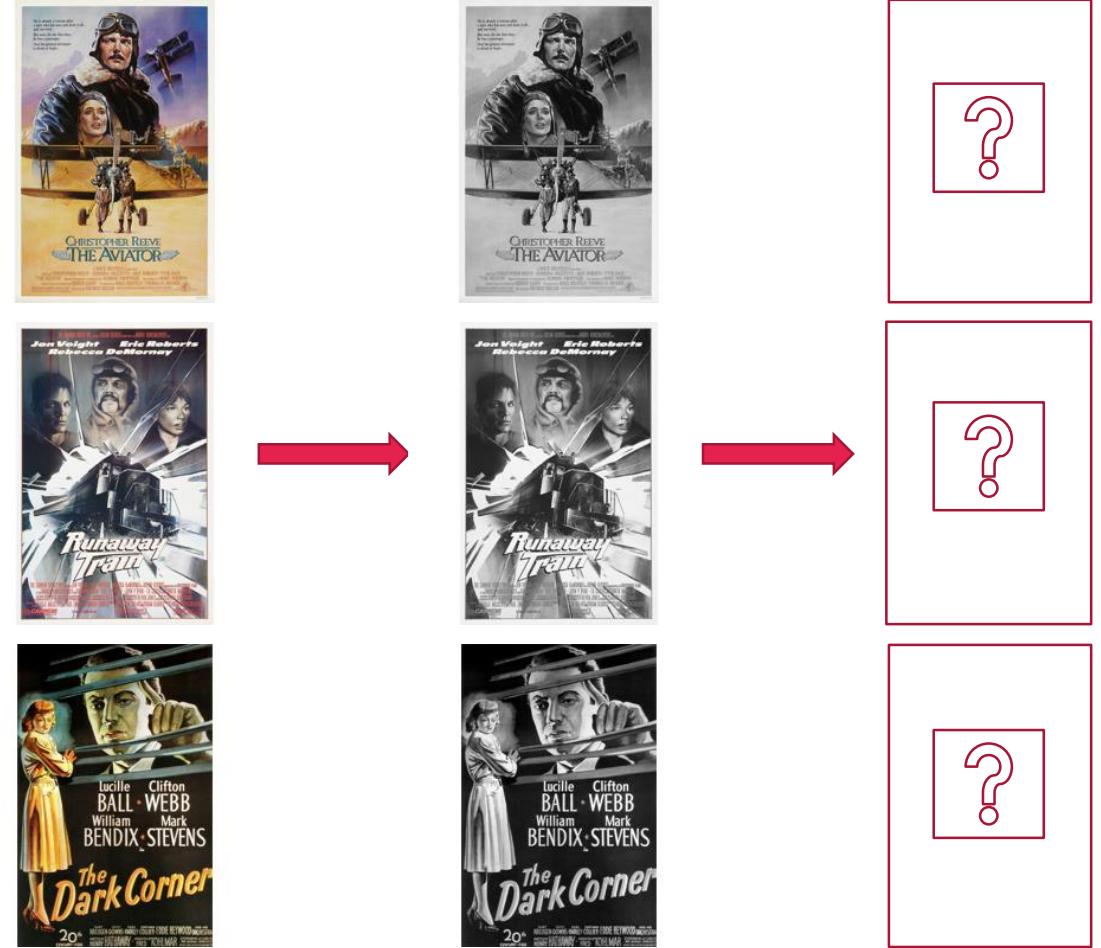
AUTOMATIC RECOLORIZATION OF MOVIE POSTERS

by Alex Meistrenko

Student Center for Professional Development
Stanford University
alexmeis@stanford.edu

Motivation:

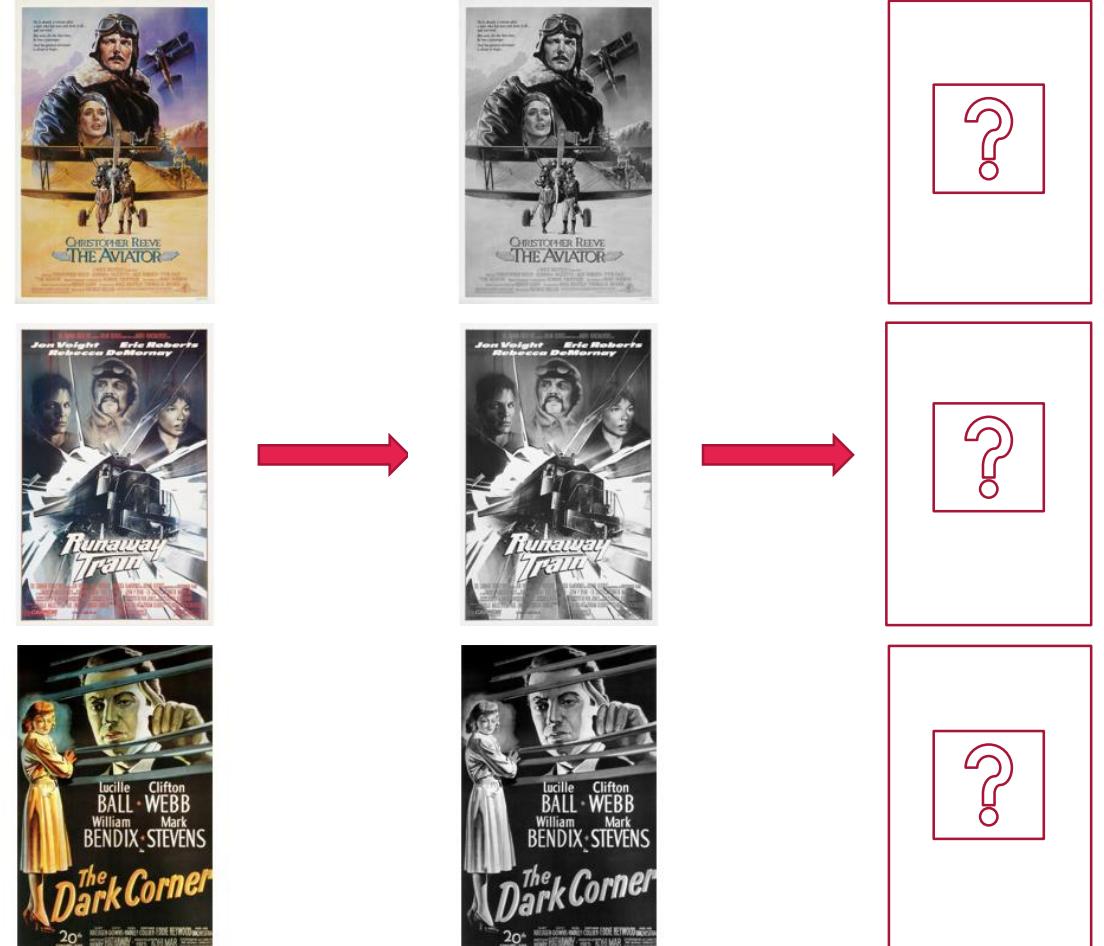
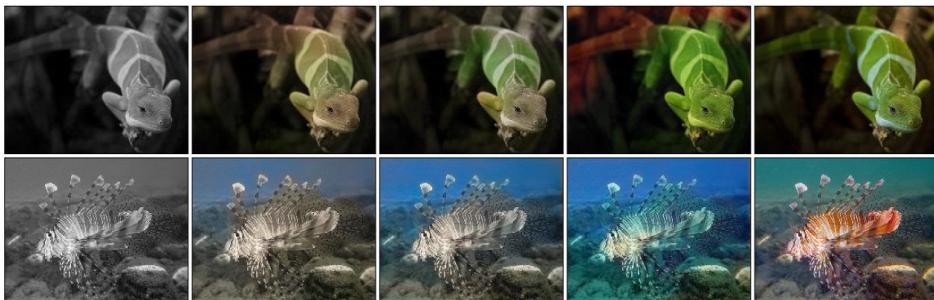
- colorization of grey-scaled images
- multimodal color classification



Motivation:

- colorization of grey-scaled images
- multimodal color classification

**What would happen,
when we apply colorization methods
to movie posters?**

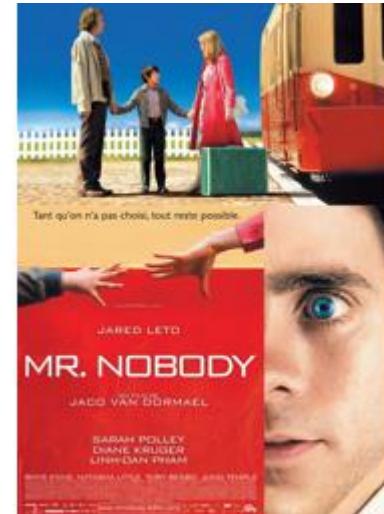


dataset:

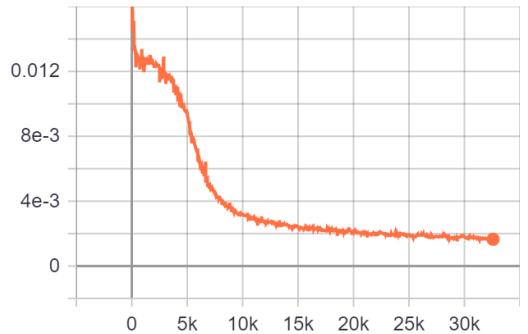
- movie posters from the time-period 2000-2020
 - 9K images for training, 1K for validation and 800 for test
- + 500 old movie posters from 1940-1990

extended dataset:

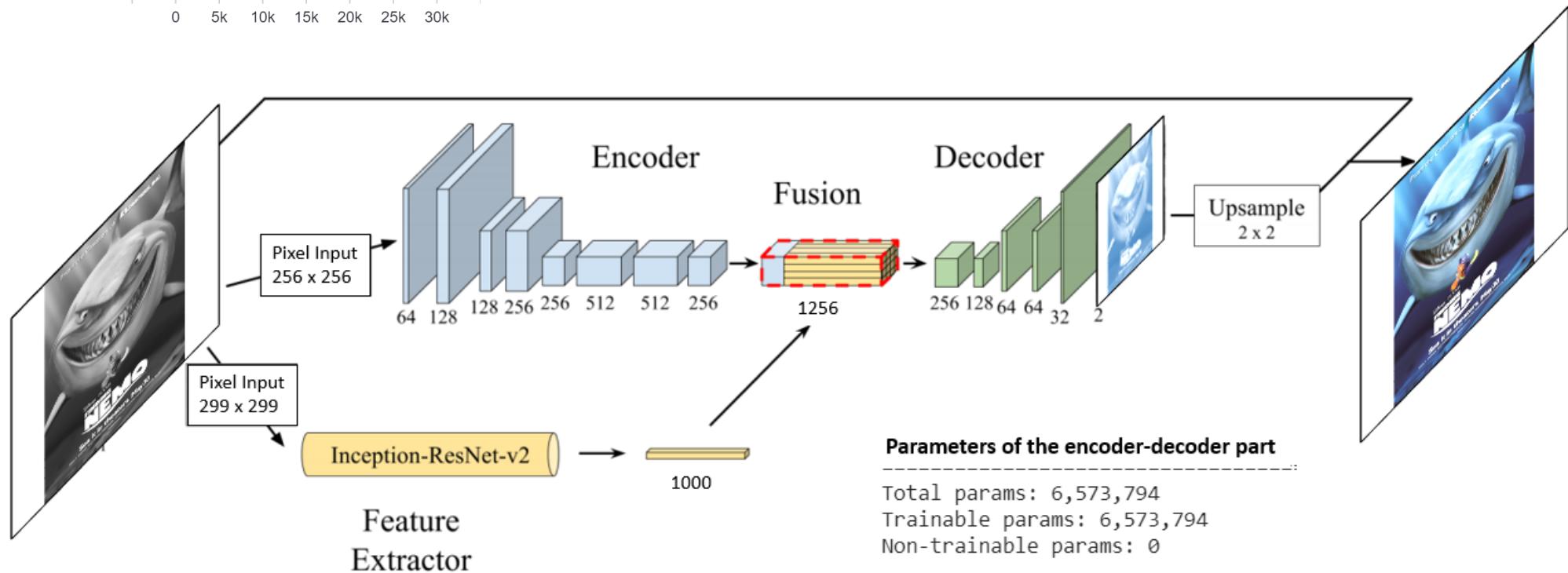
- movie posters from the time-period 2000-2020
 - 40K images (without validation/test set)



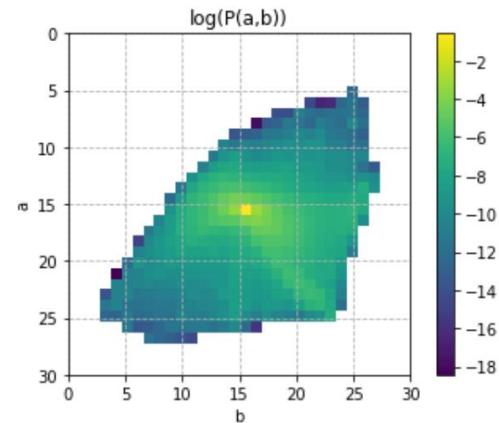
batch_loss



$$L_{2,ab} \left(X, \hat{X} \right) = \frac{1}{2HW} \sum_{k \in \{a,b\}} \sum_{h=1}^H \sum_{w=1}^W \left(X_{h,w,k} - \hat{X}_{h,w,k} \right)^2$$

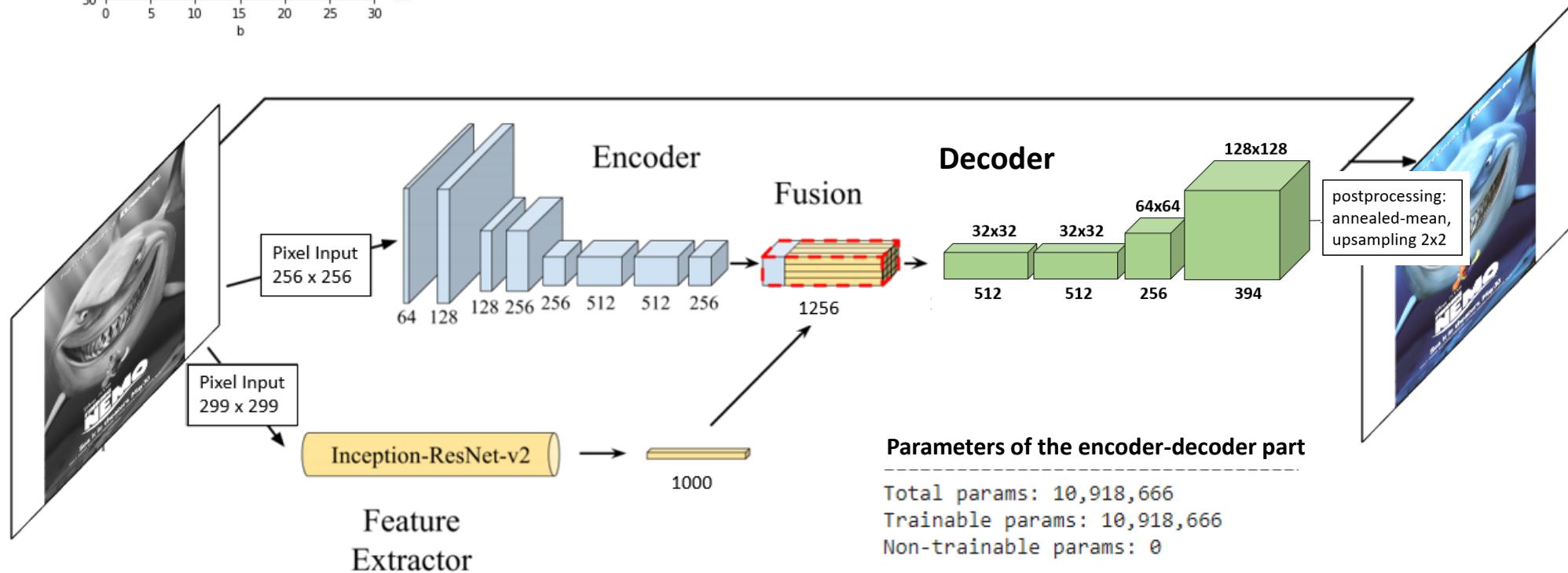


$$L_{CE,q} (Z, \hat{Z}) = - \sum_{h,w} v(Z_{h,w}) \sum_q Z_{h,w,q} \log \hat{Z}_{h,w,q}$$



$$v(Z_{h,w}) := w_{q^*}, \quad q^* = \arg \max_q Z_{h,w,q}, \quad w \sim \left(f_q + \frac{1}{Q} \right)^{-1}, \quad \mathbb{E}[w] = \sum_q f_q w_q = 1$$

$$\hat{X}_{h,w,(a,b)} = \mathcal{H}(\mathbb{E}[f_T(Z_{h,w})]), \quad f_T(z) := \frac{\exp(\log(z)/T)}{\sum_q \exp(\log(z_q)/T)}, \quad T = 0.38$$



metric	regression (trained on 9K images)	classification (trained on 9K images)
$L_{2,ab}$ (averaged)	239.99	282.50
number color classes	250	319
1-class accuracy	21.0%	23.0%
2-class accuracy	42.8%	44.8%
0-class accuracy with white padding	51.5%	52.4%

input



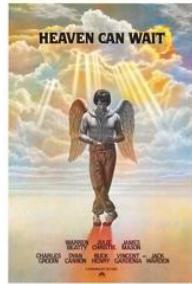
regression
training on
9K images



classification
training on
9K images



regression
training on
40K images



ground
truth



- both models show a decent performance in recoloring old movie posters
- for classification model the results are very colorful and brighter than the original images
- regression model trained on 40K images shows a colorful and more consistent colorization



metric	regression (trained on 9K images)	classification (trained on 9K images)
$L_{2,ab}$ (averaged)	239.99	282.50
number color classes	250	319
1-class accuracy	21.0%	23.0%
2-class accuracy	42.8%	44.8%
0-class accuracy with white padding	51.5%	52.4%

input



regression
training on
9K images



classification
training on
9K images



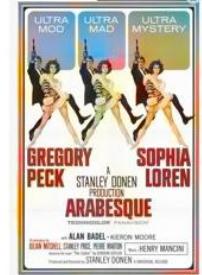
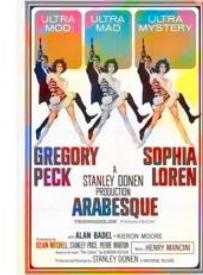
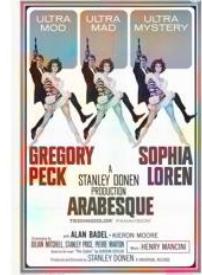
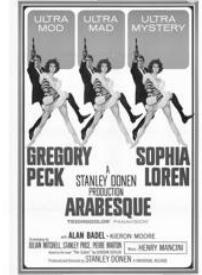
regression
training on
40K images



ground
truth

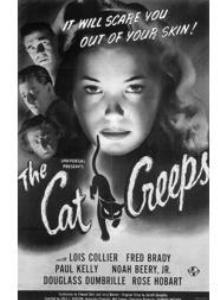
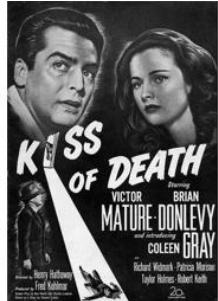


- both models show a decent performance in recoloring old movie posters
- for classification model the results are very colorful and brighter than the original images
- regression model trained on 40K images shows a colorful and more consistent colorization

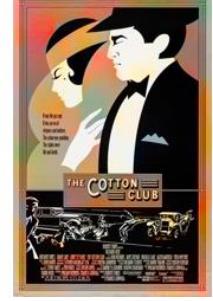


metric	regression (trained on 9K images)	classification (trained on 9K images)
$L_{2,ab}$ (averaged)	239.99	282.50
number color classes	250	319
1-class accuracy	21.0%	23.0%
2-class accuracy	42.8%	44.8%
0-class accuracy with white padding	51.5%	52.4%

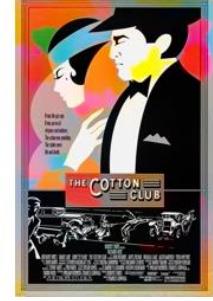
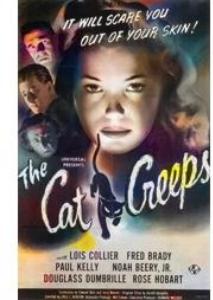
input



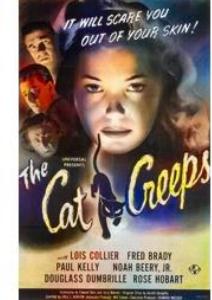
regression
training on
9K images



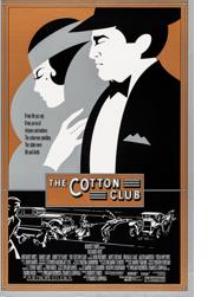
classification
training on
9K images



regression
training on
40K images



ground
truth



- both models show a decent performance in recoloring old movie posters
- for classification model the results are very colorful and brighter than the original images
- regression model trained on 40K images shows a colorful and more consistent colorization

Conclusion and Outlook

CNN encoder-decoder architecture for recoloring old movie posters

- regression model shows rather less saturated results
 - more training examples are required (40K performs much better)
- classification model shows colorful results
 - costly to train due to a large output volume

Possible development directions:

- training of the classification model on the larger dataset
- models with U-Net architecture (skipping connections)
- more artistic recolorization and style transfer with GANs