# DATASETS

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Path** | **N** | **Desc** |
| Relighted Cube+ complex segmentation | project/dataset\_relighted | 13950 | Corrected Cube+ images relighted by finding edges and contours, combination of saturated and non-saturated images |
| Combined dataset | project/dataset\_combined | 3064/2589[[1]](#footnote-1) | Corrected Cube+ images relighted using all tried segmentation models (convex hull, threshold, erosion, edges…) which contains saturated images |
| Adaptive threshold dataset | project/complex6 | 1030 | Corrected Cube+ images relighted using adaptive threshold[[2]](#footnote-2) segmentation which contains saturated images |
| Lab - projector | project/lab/combined[\_tiff] | 248 | Images taken of objects illuminated by one projector projecting an image simulating two illuminants |
| Lab - projector+ambient | project/lab/realworld[\_tiff] | 10 | Images taken of scenes where one illuminant is ambient, and one is generated by a projector |
| Crf - laboratory | project/dataset\_crf/lab | 58 | Images of objects in a dark environment illuminated by two light sources, small difference between illuminants |
| Crf - realworld | project/dataset\_crf/realworld | 20 | Images of real-world scenes with one artificial illuminant, has many unbalanced masks |
| Test | project/dataset\_test | 70 | Test images combined from all datasets to evaluate model performance |

**Dataset directory structure:**

|\_\_ images - Images used as the input for the regression model and model testing

|\_\_ gt - The groundtruth infoirmation about the illumination for each pixel

|

|\_\_ img\_corrected\_1 - Images corrected for one illuminant, used to train segmentation models

|\_\_ gt\_mask - The mask used for segmentation[[3]](#footnote-3)

# RESULTS

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Segmentation | | | |  |  |  |
| Model | Dataset | | Dice | | | |
| Train | Test | Mean | Std | Max (id) | Min (id) |
| U-Net efficientnet b0 | Combined | [V1@dataset\_test](mailto:V1@dataset_test) | 0.570 | 0.177 | 0.809 (9) | 0.003 (37) |
| Adaptive thresh | [V1@dataset\_test](mailto:V1@dataset_test) | 0.587 | 0.104 | 0.803 (9) | 0.292 (64) |
| Custom U-Net | Adaptive thresh | [V1@dataset\_test](mailto:V1@dataset_test) | 0.357 | 0.200 | 0.797 (5) | 0.101 (10) |
| U-Net efficientnet b2 | Desaturated | [V1@dataset\_test](mailto:V1@dataset_test) | 0.586 | 0.158 | 0.807 (2) | 0.005 (66) |

Table 2. Results of the segmentation model on the test dataset. The desaturated dataset refers to the original generated dataset with ~ 9000 images with only desaturated illuminants.

Discussion:

* Custom architecture was to simple to achieve any sort of competitive results
* Little difference between the EfficientNet U-Nets, however the B2 architecture is still being trained and could improve
* There isn’t much difference in between the median and the mean values, so only mean is shown

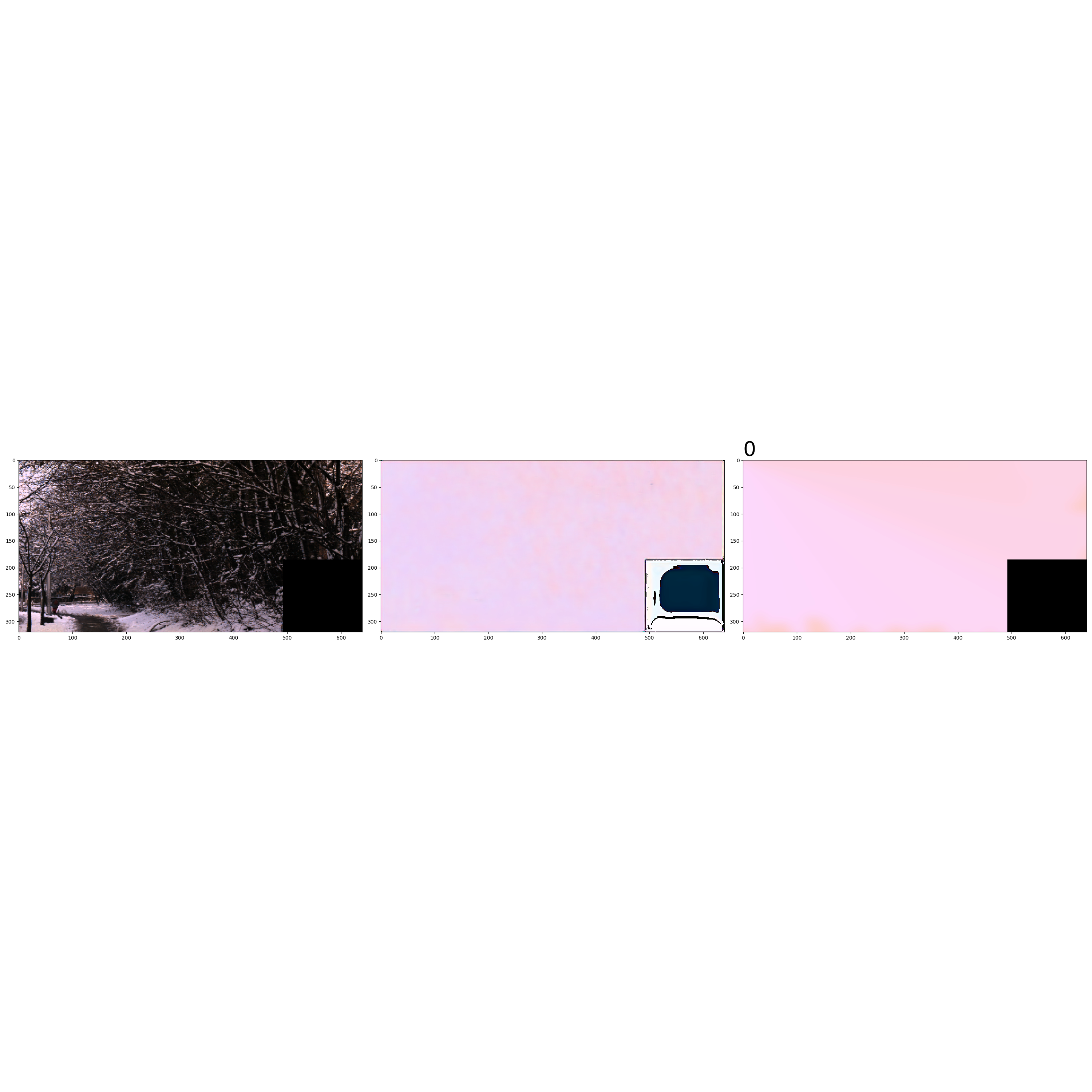
|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Regression | | | | | | |  |  |
| Model | | Dataset | | Angular error per-pixel | | | | | |
| Train | Test | Mean | Std | Med. | Trimean | Min (id) | Max (id) |
| U-Net Efficientet b2 | | Desaturated | [V1@dataset\_test](mailto:V1@dataset_test) | 11.56 | 10.37 | 5.125 | 9.353 | 1.094 (15) | 26.50 (68) |
| Crf laboratory | 25.52 | 0.779 | 25.51 | 25.51 | 24.14 (47) | 25.82 (1) |
| Crf realworld | 26.24 | 0.953 | 26.21 | 26.17 | 24.45 (5) | 28.73 (9) |
| Generated test set | 4.561 | 4.543 | 2.884 | 3.168 | 1.032 (15) | 22.06 (22) |
| Cube 12500 (desaturated + saturated) | [V1@dataset\_test](mailto:V1@dataset_test) | 10.232 | 9.081 | 4.134 | 8.261 | 1.488 (1) | 25.75 (64) |
| Crf laboratory | 22.63 | 1.233 | 22.41 | 22.53 | 20.84 (23) | 25.83 (1) |
| Crf realworld | 25.60 | 1.025 | 25.86 | 25.95 | 22.95 (5) | 26.80 (9) |
| Generated test set | 3.840 | 2.327 | 3.312 | 3.262 | 1.488 (1) | 13.22 (5) |
| Cube 15000 (desaturated + saturated) | [V1@dataset\_test](mailto:V1@dataset_test) | 10.57 | 10.52 | 3.835 | 8.686 | 1.094 (0) | 27.18 (64) |
| Crf laboratory | 25.74 | 0.959 | 25.66 | 25.69 | 24.46 (23) | 28.25 (1) |
| Crf realworld | 26.69 | 1.007 | 26.96 | 27.03 | 24.16 (5) | 28.24 (12) |
|  | | Generated test set | 3.021 | 1.351 | 2.897 | 2.901 | 1.094 (0) | 8.587 (5) |

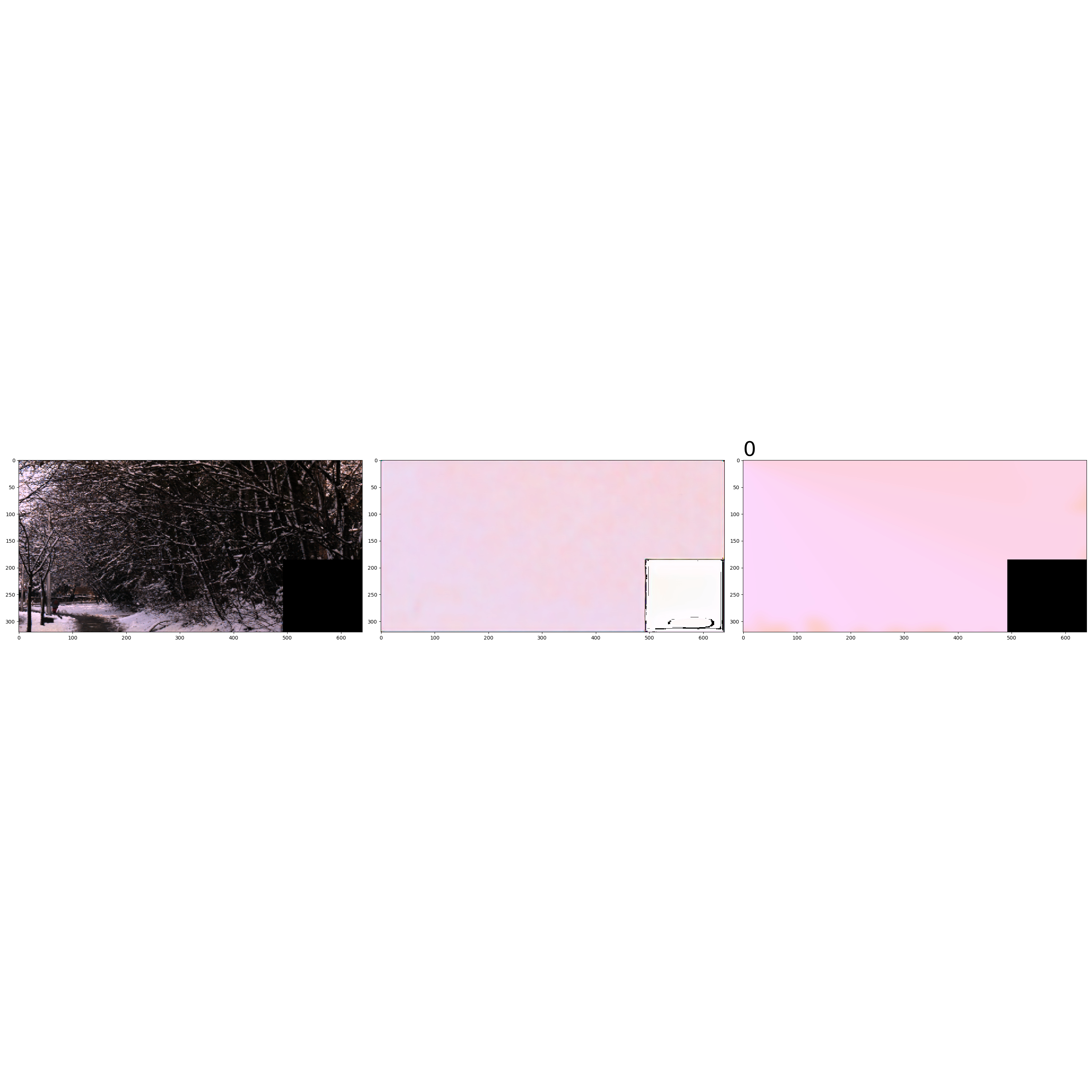
Table 3. Results of the regression models trained on different datasets. Experiments were conducted on the test dataset and the crf[[4]](#footnote-4) dataset

Discussion:

* High mean value and low trimean and median values indicate that there are numerous images for which the regression works well and a few for which it does not
* That is why additional experiments were conducted on the crf dataset, and the results show that the models perform bad on those images, which suggests overfitting on the train dataset
* It was noted that the saturated colors were problematic for these models (Appendix)

# Appendix

Figure 1: Output of the first regression model on the image with index 0 from the generated dataset, good performance of the model

 Figure 2: Output of the second regression model on the image with index 0 from the generated dataset, good performance of the model

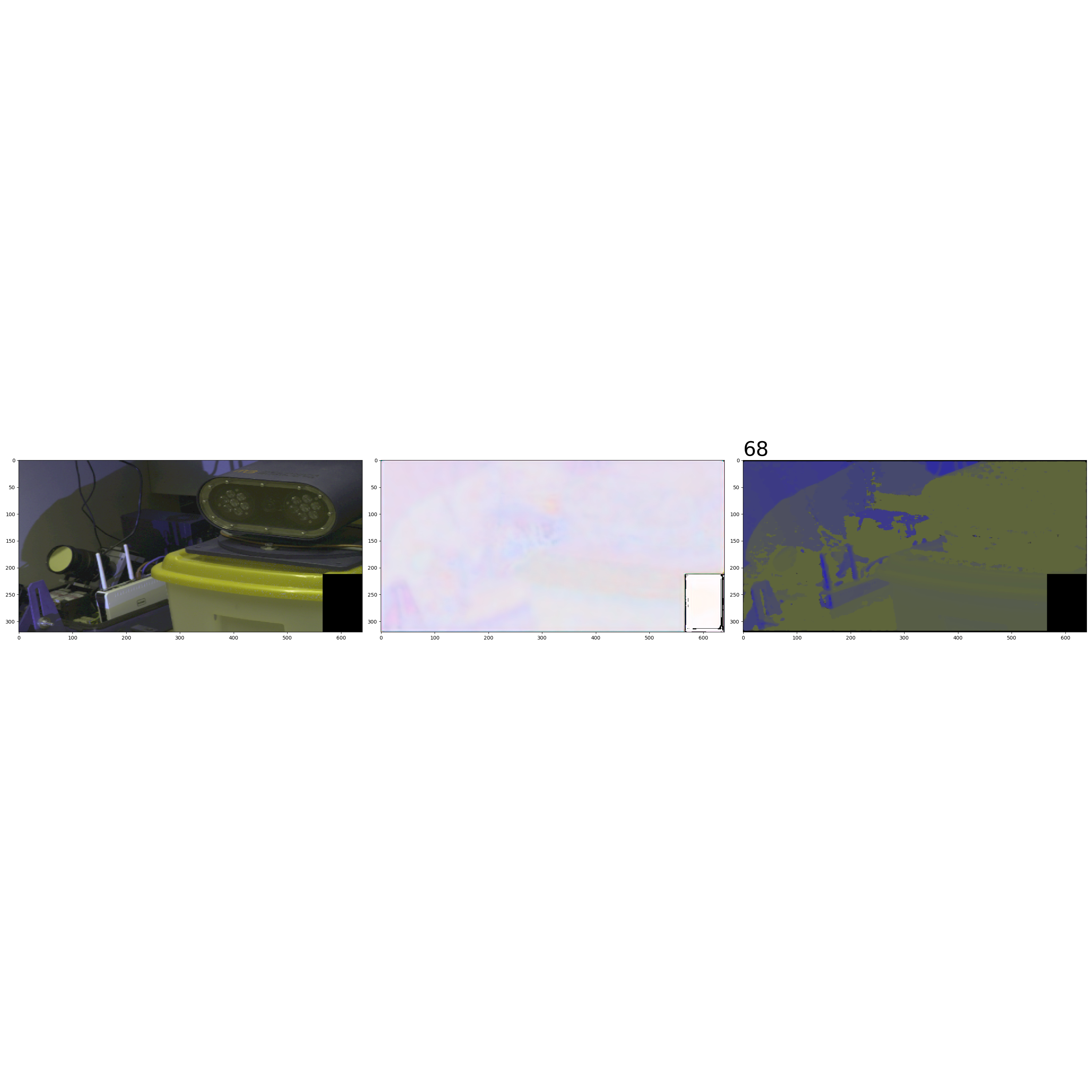


Figure 3: Output of the first regression model on the image with index 68 from the crf realworld dataset, poor performance of the model

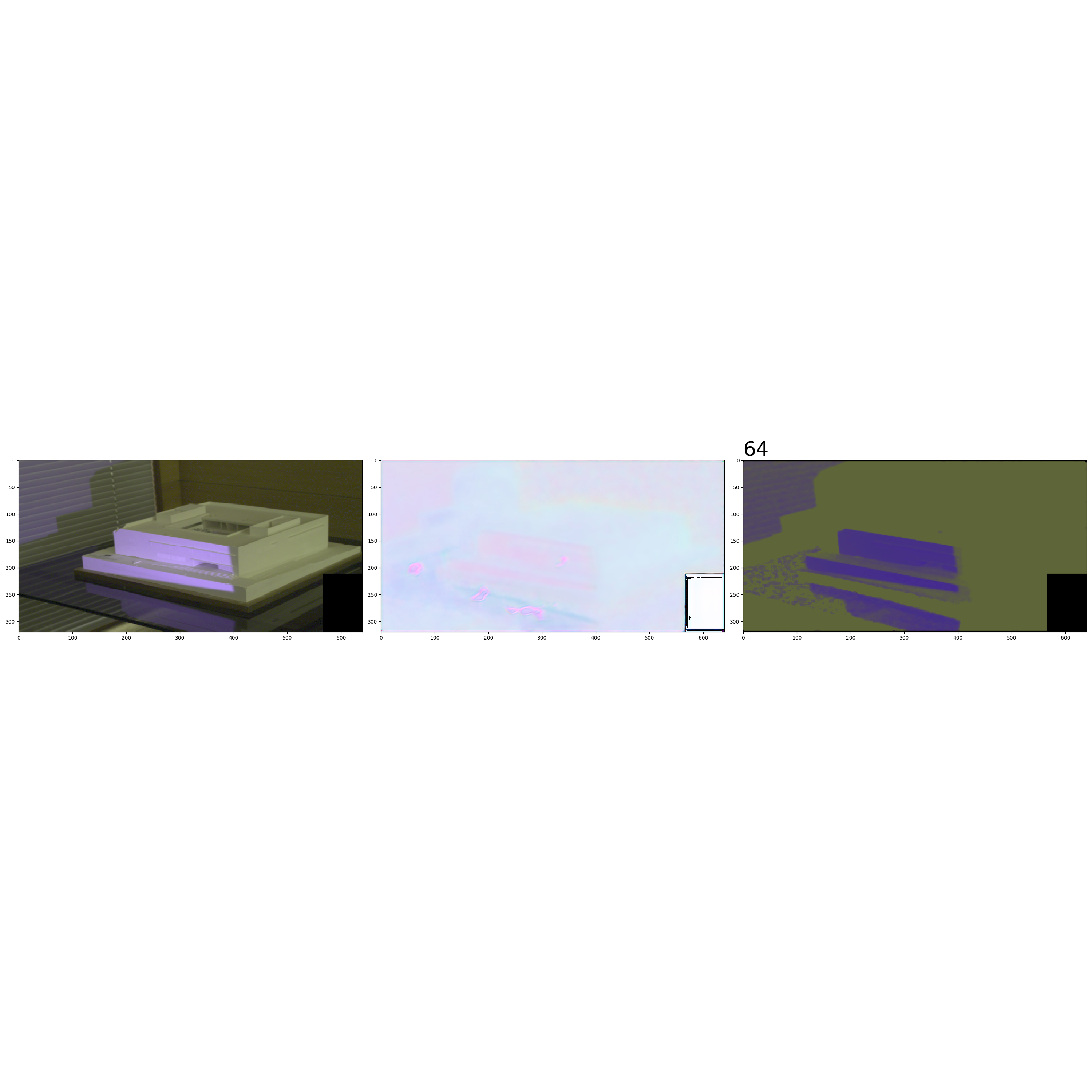
****

Figure 4: Output of the third regression model on the image with index 64 from the crf realworld dataset, poor performance of the model

1. Regression/Segmentation [↑](#footnote-ref-1)
2. Adaptive threshold – threshold for detecting contours was acquired taking the mean luminance and adding a random shift vector [↑](#footnote-ref-2)
3. Obtained by clustering illuminant values in groundtruths and comparing each pixel to one of the centers to determine to which illuminant it is closer [↑](#footnote-ref-3)
4. Shida Beigpour, Christian Riess, Joost Weijer, i Elli Angelopoulou. Multi-illuminant estimation with conditional random ﬁelds. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, 01 2014. doi: 10.1109/TIP. 2013.2286327. [↑](#footnote-ref-4)