Internship Technical Report - Report 3

Jessica Klebe
j.s.klebe@gmx.de

May 20, 2021

1 Pipeline

The aim was to determine the image which has - according to the model - the highest probablity to belong to one specific class. For each image of the test set the values of the pixels that are classified as the same class are summarized and divided by the number of these pixels. Finally, for each class the image with the highest values is determined. Fig. 1 depicts the respective images.

2 LinkNet34

The Unet model from the last report has been substituted by the LinkNet34 model where the early stopping method is based on the IoU value. The results are depicted in Tab. 1 where the values of the Unet are compared with the ones from the LinkNet34. It becomes apparent that the values are very similar and both networks perform similarly well.

3 Postprocessing

In order to improve the outcome the predictions have been augmented according to the following algorithm: First, each pixel of the class that is predicted most (except for background) is magnified to a circle with a radius of 11 pixels. Second, each pixel predicted as background is also enlarged to a circle with a radius of 11 pixels. This closes holes. The same procedure is repeated, however, first the background is magnified, and afterwards the class. This decreases clutter.

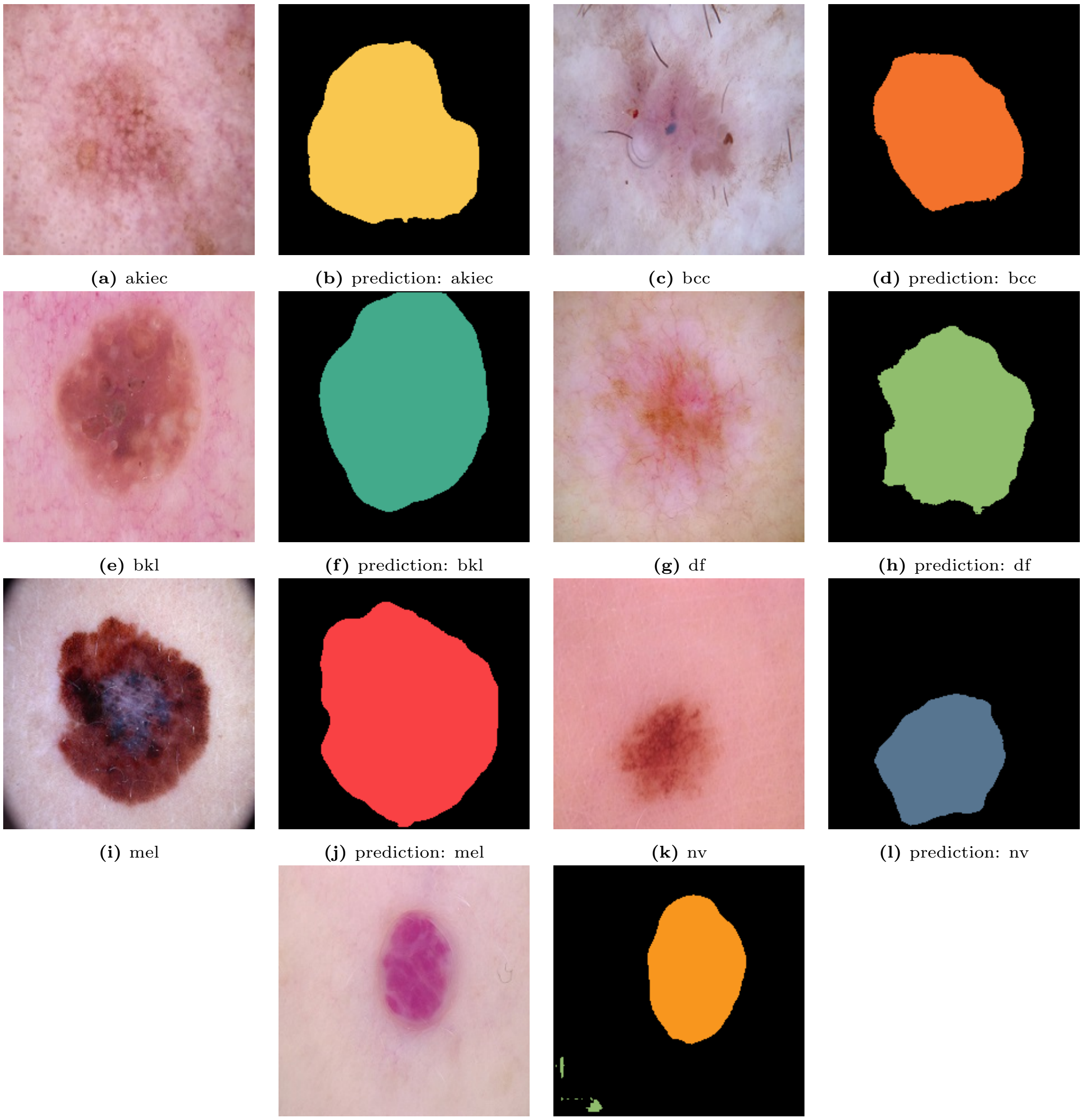
Two examples are given in Fig. 2. In the first example a small hole is closed, in the second one clutter is removed.

In Tab. 2 the performance of LinkNet34 with and without postprocessing is compared. It shows that except for PixSim on the test set the performance is slightly better when postprocessing is applied, however, overall the performance is comparably similar.

Fig. 3 depicts the relative number of images where the lesion has been determined correctly up to a certain percentage in dependence on this percentage value. In other words, at the value 0.9 on the x-axis the value on the y-axis equals No_lesion_90. Fig. 3 shows the results for the validation and test sets each with and without postprocessing. The results for both the validation and the test set illustrate that postprocessing improves the performance for predictions of well-predicted lesions while, if the lesion is detected less than 40 % correctly, the performance improves if no postprocessing is used. Hence, postprocessing improves the performance in cases where the results are already good while it slightly magnifies the problem of low-scoring images. The latter is expected because postprocessing is prone to overwrite pixels of a class with the dominant class. Particularly for low-scoring images the model might predict several classes such this effect might be seen.

4 XGrad-CAM-Analysis

An XGrad-CAM Analysis [1] has been conducted. In Fig. 4 and 5 an example is given for each class where the Unet's prediction without postprocessing, XGrad-CAMs analysis and mask are compared. For



(a) akiec **(b)** prediction: akiec **(c)** bcc **(d)** prediction: bcc

(e) bkl **(f)** prediction: bkl **(g)** df **(h)** prediction: df

(i) mel **(j)** prediction: mel **(k)** nv **(l)** prediction: nv

(m) vasc **(n)** prediction: vasc

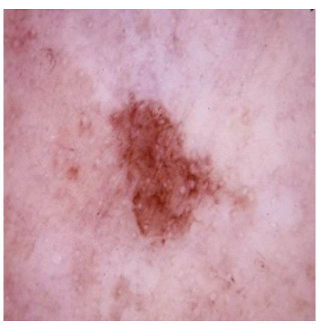
Figure 1: Images which have - according to the model - the highest probability to belong to the respective class.

(a) Example 1: image

(b) Example 1: prediction

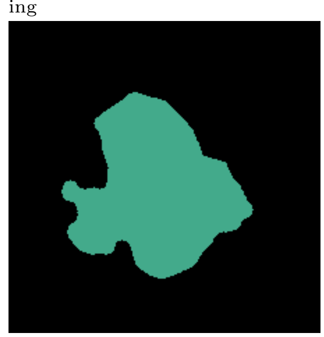
(c) Example 1: after postprocessing

(d) Example 1: mask



(e) Example 2: image

(f) Example 2: prediction



(g) Example 2: after postprocessing

(h) Example 2: mask

Figure 2: Two examples for postprocessing.

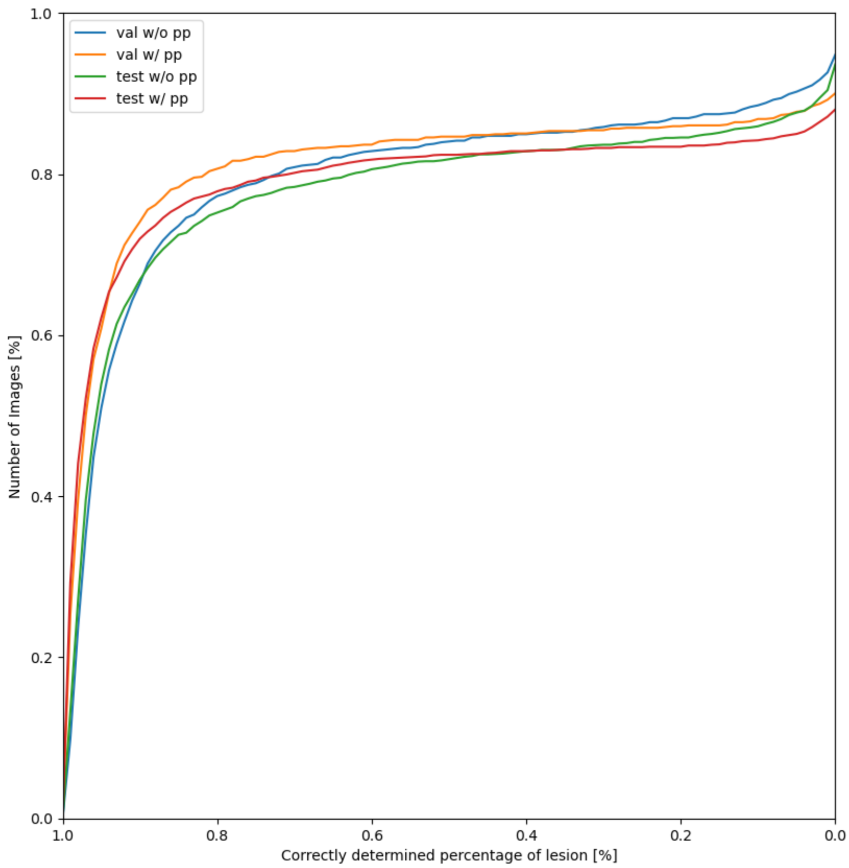


Figure 3: The relative number of images where the lesions has been determined up to a certain percentage value correctly in dependence on this percentage value. The values for the validation and test sets with and without postprocessing are compared. The model used is LinkNet34 where the early stopping method has been optimized based on the IoU score.

Table 1: Scores and auxiliary values for the validation and test sets of the Unet and LinkNet34 where the network has been trained with an early stopping condition based on the IoU score.

Score or Auxiliary	Val Unet	Test Unet	Val LinkNet	Test LinkNet
PixSim	0.8887	0.8772	0.8851	0.8732
PixSim_black	0.6469	0.6339	0.6469	0.6339
Jaccard	0.2141	0.1032	0.2043	0.1028
IoU	0.5961	0.5103	0.5488	0.4929
N_lesion_90	0.6580	0.6479	0.6640	0.6688
N_lesion_50	0.8285	0.8015	0.8404	0.8185
N_90	0.7986	0.7461	0.7647	0.7501
N_90_black	0.1176	0.1187	0.1176	0.1187

Table 2: Scores and auxiliary values for the validation and test sets where the network has been trained with an early stopping condition based on IoU. The model used is the LinkNet34 model. Comparison of the performance without (left side) and with (right side) postprocessing.

Score or Auxiliary	Val	Test	Val (postprocessing)	Test (postprocessing)
PixSim	0.8851	0.8732	0.8872	0.8609
Jaccard	0.2043	0.1028	0.2074	0.1037
IoU	0.5488	0.4929	0.5642	0.5082
N_lesion_90	0.6640	0.6688	0.7408	0.7197
N_lesion_50	0.8404	0.8185	0.8465	0.8239

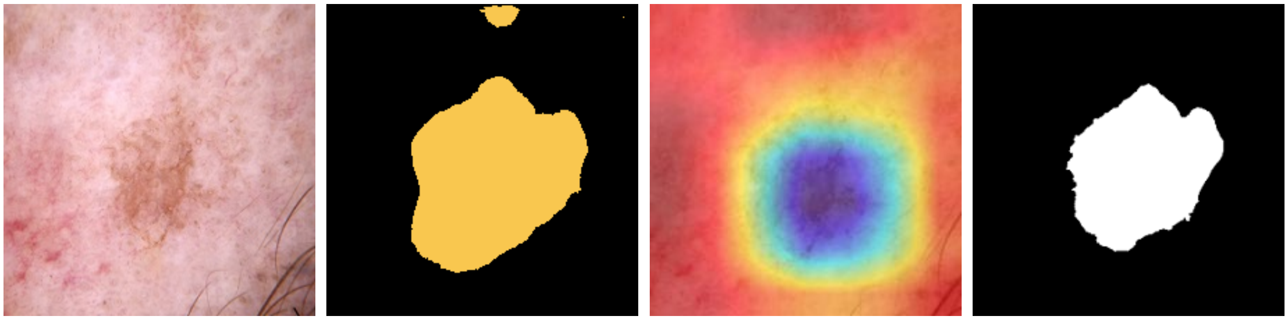
the class bcc the model does not recognize the lesion but mistakes the background for a lesion while the XGrad-CAM method does recognize the lesion in this particular case. In all other depicted cases the model and the XGrad-CAM method detect the lesion similiarly well.

5 Jaccard Loss

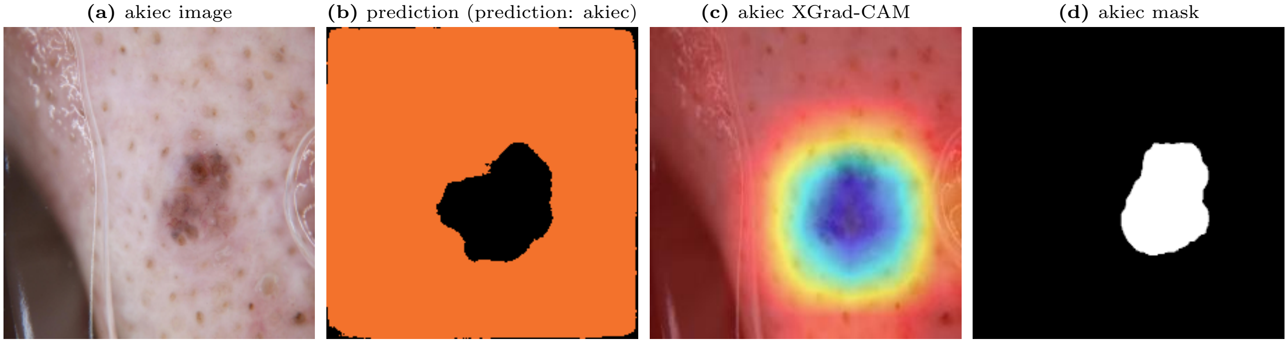
The model described in the previous report has been trainded now using the Jaccard Loss function from https://github.com/ternaus/robot-surgery-segmentation/blob/master/loss.py. However, it has not increased the model's performance.

References

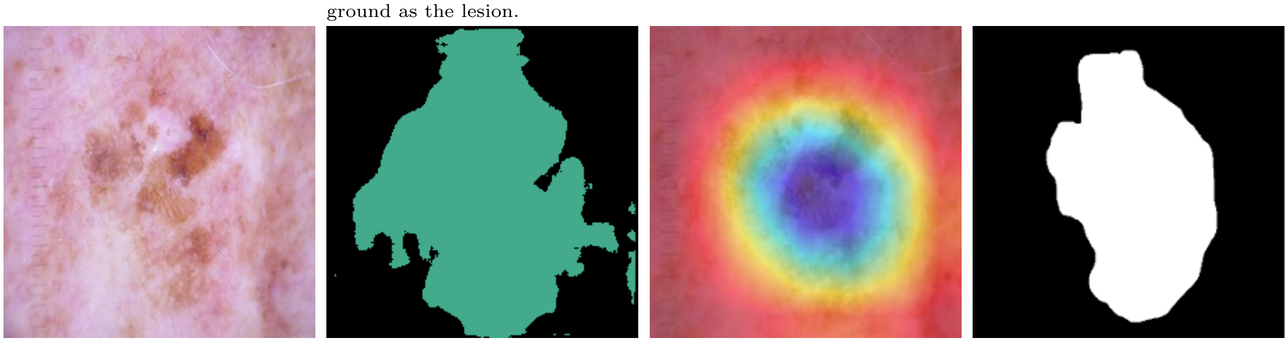
[1] Ruigang Fu, Qingyong Hu, Xiaohu Dong, Yulan Guo, Yinghui Gao, and Biao Li. Axiom-based grad-cam: Towards accurate visualization and explanation of cnns. 2020.



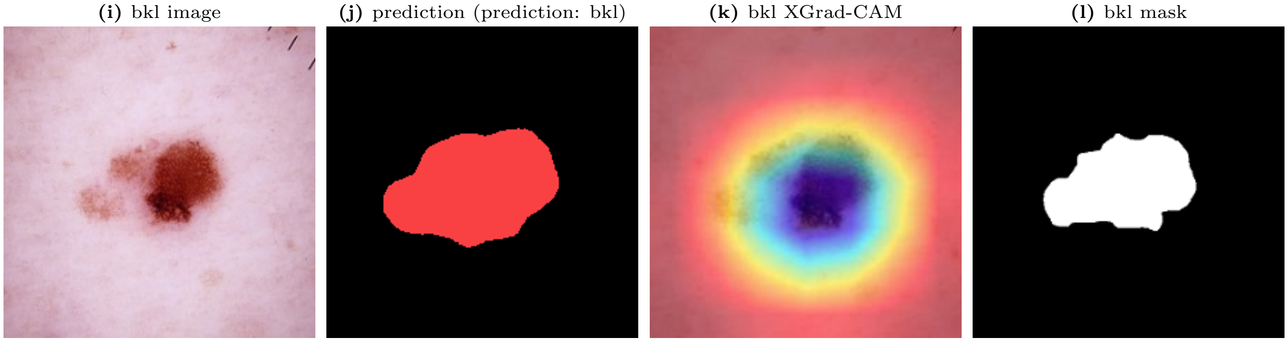
(a) akiec image **(b)** prediction (prediction: akiec) **(c)** akiec XGrad-CAM **(d)** akiec mask



(e) bcc image **(f)** prediction (prediction: bcc). Attention: This is an example where the net does not recognize the lesion but mistakes the background as the lesion. **(g)** bcc XGrad-CAM **(h)** bcc mask

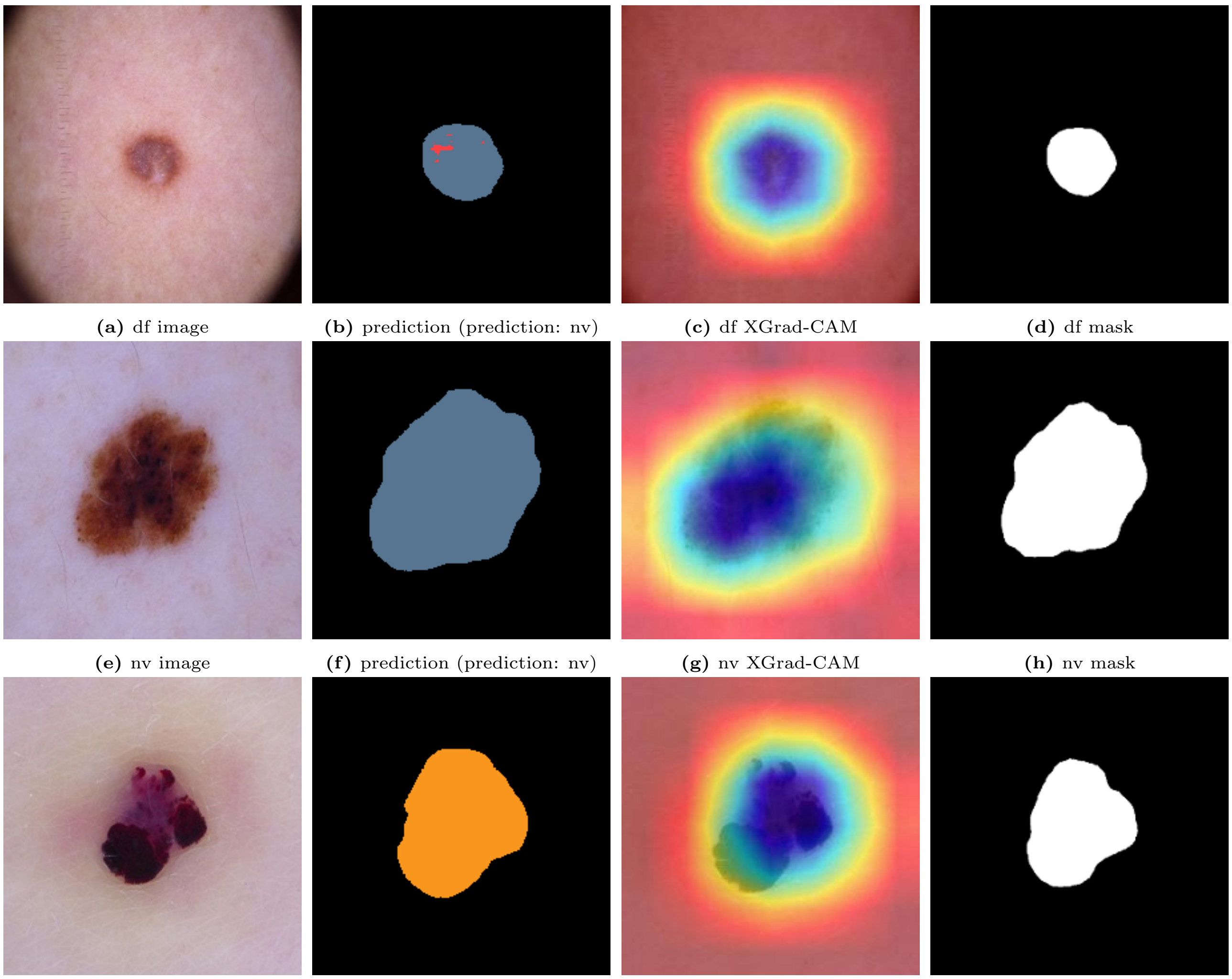


(i) bkl image **(j)** prediction (prediction: bkl) **(k)** bkl XGrad-CAM **(l)** bkl mask



(m) mel image **(n)** prediction (prediction: mel) **(o)** mel XGrad-CAM **(p)** mel mask

Figure 4: An example of the Unet's prediction and XGrad-CAM prediction for each class - part 1.



(a) df image **(b)** prediction (prediction: nv) **(c)** df XGrad-CAM **(d)** df mask

(e) nv image **(f)** prediction (prediction: nv) **(g)** nv XGrad-CAM **(h)** nv mask

(i) vasc image **(j)** prediction (prediction: vasc) **(k)** vasc XGrad-CAM **(l)** vasc mask

Figure 5: An example of the Unet's prediction and XGrad-CAM prediction for each class - part 2.