Link to repository: https://github.com/emgeerthsen/2026_PDS_Rabbits.git

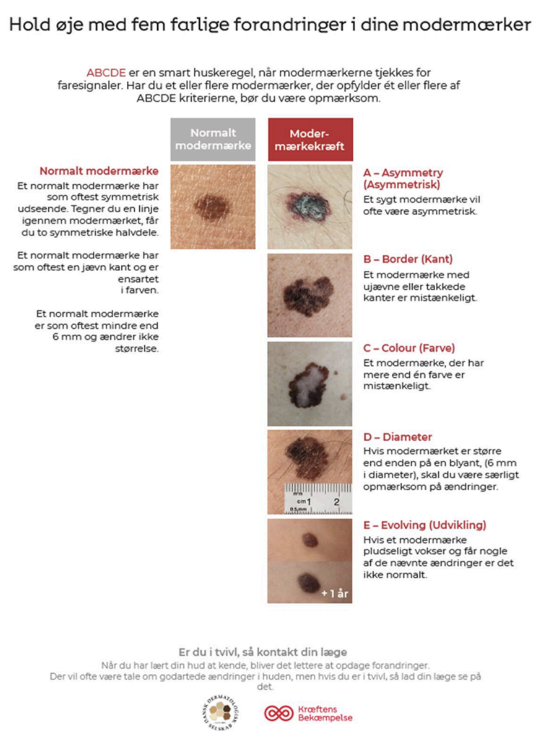
Contributors:

Name	*itu mail*	*Github username*
Alexandra Istrate	istr@itu.dk	BirbTamer
Emilie Lass Geerthsen	*emge@itu.dk*	emgeerthsen
Mihnea-Andrei Simion	misi@itu.dk	HeavenlyDemon47
Sidra Al-Montaha Tasharoki	siat@itu.dk	sidra-Al-Montaha

Skin cancer is the 17th most common cancer type worldwide, and is the 14th most common cancer type for both men and women globally.[1]
There are different subtypes of skin cancer, but it can be divided into 2 main groups: melanoma and non-melanoma. Melanoma is a more serious type of skin cancer.[2]

The most common non-melanoma cancers are:

- basal cell carcinoma. This is the most common form of skin cancer and develops in basal cells in the epidermis layer of the skin.
- squamous cell carcinoma. This is the second-most common form of skin cancer and develops in squamous cells in the epidermis layer of the skin.



When determining cancer in skin lesions, you can use the ABCDE-rule[3], where you look for asymmetry, border, colour, diameter and evolving, but there is even more features to look for:[4]

- dots
- globules
- lines
- network structures
- regression structures
- Vessels

[1] https://www.wcrf.org/preventing-cancer/cancer-statistics/skin-cancer-statistics/
[2] https://www.wcrf.org/preventing-cancer/cancer-types/skin-cancer/
[3] https://www.aad.org/public/diseases/skin-cancer/find/at-risk/abcdes
[4] https://pmc.ncbi.nlm.nih.gov/articles/PMC9892985/

However this project focuses on ABC, since D, the distance at which the photos are taken is varying and E, we don't have access to more than one picture per lesion, meaning we don’t know if the lesion has evolved. The dataset that was given has 6 skin lesion types.

Humans diagnose cancer by focusing on medically meaningful features based on training and context. In contrast, algorithms learn statistical patterns from data and can latch onto visual shortcuts that have nothing to do with real disease signals.

Small changes in pixel color can lead to incorrect computer-based cancer diagnoses because digital systems rely on precise color values and boundaries. Variations in staining, scanner settings, image compression, or preprocessing can slightly alter pixel intensity. Minor shifts may change how tissue is classified, potentially turning a benign region into a suspicious one, or masking a real tumor.

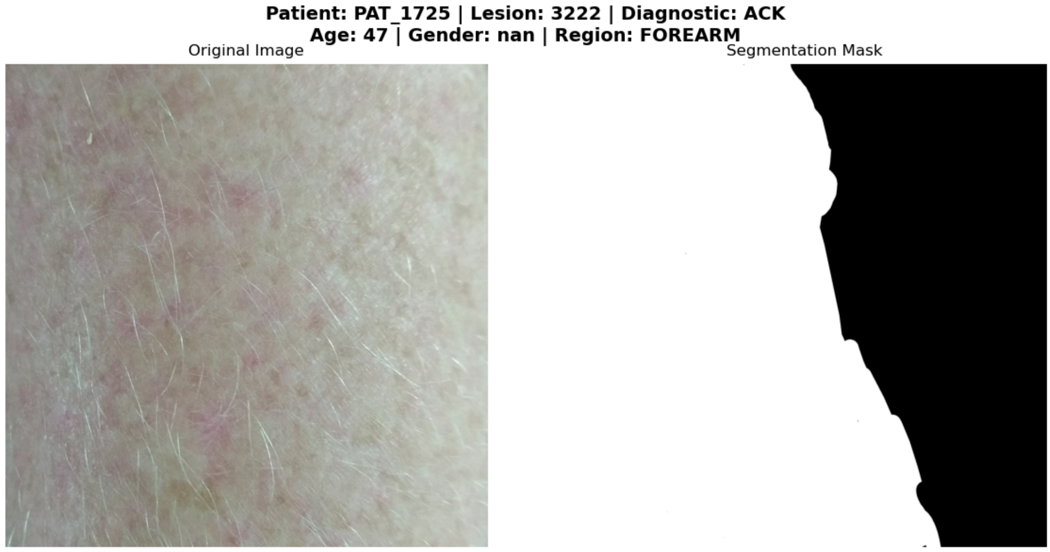
Cat(Dog) Dog(Cat)

In cancer diagnosis, this means an algorithm might learn to use background noise, scanner artifacts, or dataset biases as cues for “cancer” versus “no cancer,” rather than true biological indicators. Such shortcuts can produce high accuracy in training but fail in real clinical settings.

In our image analysis, pigmented lesions, especially those with irregular borders or color variation, often resembled pen marks and created challenges for accurate classification, whereas vascular birthmarks mainly affected color-based detection due to their reddish or purplish tones.

The number of masks does not equal the number of all images in the whole dataset, such that there are more images than masks. However, the number of images in our group matches the number of masks. It is hard to determine the degree in which the masks fit the images, as many of the skin lesions do not have clear boundaries, which means that the masks are approximations of the skin lesions.

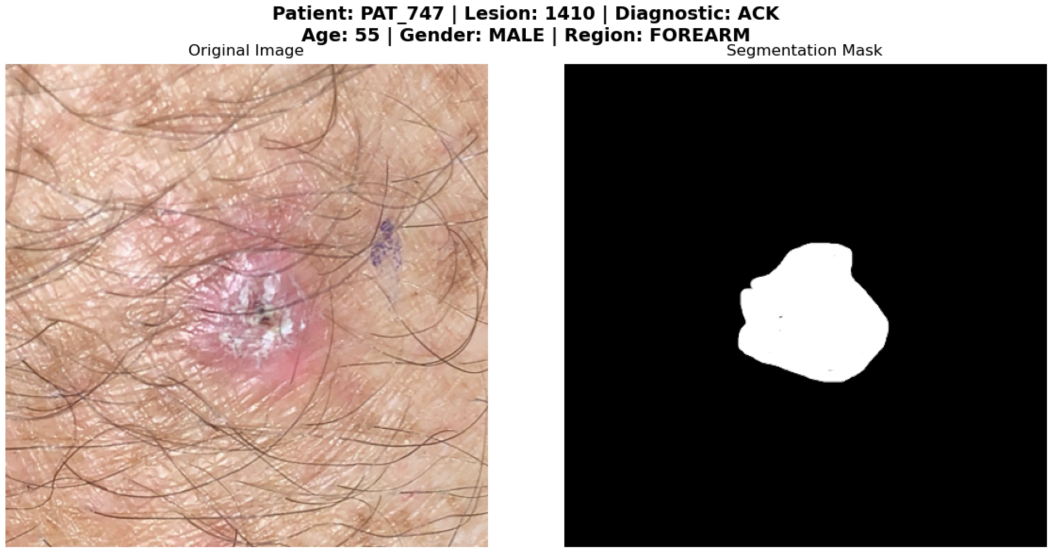
An example of a skin lesion with unclear boundaries is shown below:



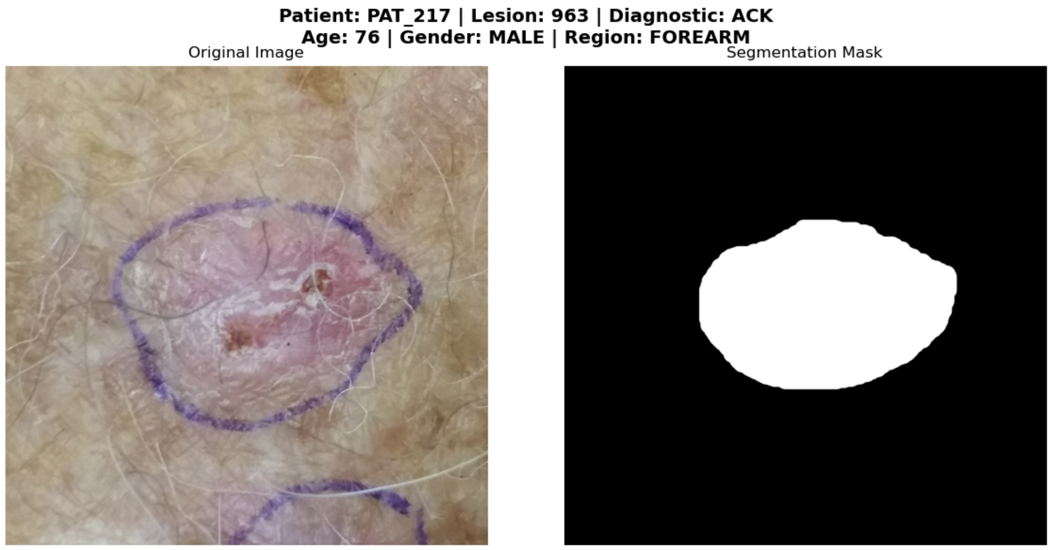
In this image, the boundaries of the skin lesion is unclear, but it covers a wide area without clear bounds, which makes comparing it to its mask difficult. In some cases, the masks clearly do not fit the images. An example of this is shown below:

When comparing this image with its mask, we can see that the skin lesion in the image is much bigger than the mask represented by the white area. A reason for this difference in this example might be that the mask is zoomed out. This is a prime example of a mask not fitting well an image.

For the following image, we observe that the presence of hair reduces the program's ability to accurately identify the mask. This occurs because the pixel colors and textures associated with hair often overlap with or obscure the features that the algorithm relies on for segmentation. As a result, the model may misinterpret hair pixels as part of the region of interest or fail to correctly distinguish the mask boundaries. The variation in hair color, thickness, and distribution further increases this difficulty, introducing noise into the image and leading to less precise mask generation. Consequently, areas covered by hair are more prone to segmentation errors and reduced detection accuracy.



In contrast, the presence of pen marks helps the program identify the mask more accurately. Pen marks typically create clear and well-defined boundaries with strong contrast relative to the surrounding skin or background, making them easier for the algorithm to detect. The consistent color intensity and sharp edges of these markings provide distinctive pixel patterns that support more reliable segmentation. As a result, the program can more effectively differentiate the region of interest from neighboring areas, leading to improved mask localization and overall detection accuracy.



For the 116 images that were assigned to our group, we annotated each of the images individually to ensure avoiding bias from being influenced by the other annotators' annotations. In this way, we have variations in annotations of each category; pen marks and amount of hair in each image.

We assessed inter-annotator agreement for the categorical variables using Cohen's Kappa statistic. The findings show a high level of agreement for pen-mark annotations, with Kappa scores between annotators ranging from 0.75 to 0.97. In comparison, the agreement for hair annotations is considerably lower, with Kappa values ranging from 0.55 to 0.77. This is because the pen annotation was easier than the amount of hair annotation. There were several disagreements: about the meaning of the amount of hair among us, whether we should annotate the hair that is at the centre or the diseased part of the skin, which led to great variations in annotating some images. Hair colour also had an impact on the bias in our annotations. Since black hair was more obvious to our eyes, these images were annotated to have a bigger amount of hair instead of the blond or white hair which was less obvious to our eyes, leading us to confuse between skin and hair, which affected both the precision and accuracy of our annotations.
Two examples of images in which the amount of hair is annotated to be between zero and three is shown below:

PAT_93_361_467

PAT_2016_4160_868

In the latter image, there is almost no hair existing at the centre of the skin lesion. However, there is so much hair around the skin lesion, which has led to a disagreement about whether the amount of hair should be annotated to be zero or higher.

Annotating pen marks had some challenges too. In some cases, it was unclear whether a blue mark was a pen mark, a blood vessel or an injury related to the disease, leading to different interpretations of the mark and hence different annotations.

The following two images are examples of this:

PAT_2006_4123_719

PAT_302_650_477

In these two images, it is unclear whether the blue marks are pen marks or injuries related to the diseased skin, which has led to variations in our annotations.