# 1 Description

| | |
|---|---|
| **Dataset name** | Breast Histopathology Images |
| **Link** | Kaggle |
| **License** | CC0: Public Domain |
| **Medical discipline** | Histopathology |
| **Medical procedure** | Microscopy |
| **Multiclass problem** | ✗ |
| **Multilabel problem** | ✗ |

Invasive Ductal Carcinoma (IDC) is the most common subtype of all breast cancers. To assign an aggressiveness grade to a whole mount sample, pathologists typically focus on the regions which contain the IDC. As a result, one of the common pre-processing steps for automatic aggressiveness grading is to delineate the exact regions of IDC inside of a whole mount slide.

The original dataset consisted of 162 whole mount slide images of Breast Cancer (BCa) specimens scanned at 40x. From that, 277,524 patches of size 50 x 50 were extracted (198,738 IDC negative and 78,786 IDC positive). Each patch's file name is of the format: uxXyYclassC.png → example 10253idx5x1351y1101class0.png . Where u is the patient ID (10253idx5), X is the x-coordinate of where this patch was cropped from, Y is the y-coordinate of where this patch was cropped from, and C indicates the class where 0 is non-IDC and 1 is IDC.

The original files are located here.

Credit the data by citing these publications:

- Publication 1
- Publication 2

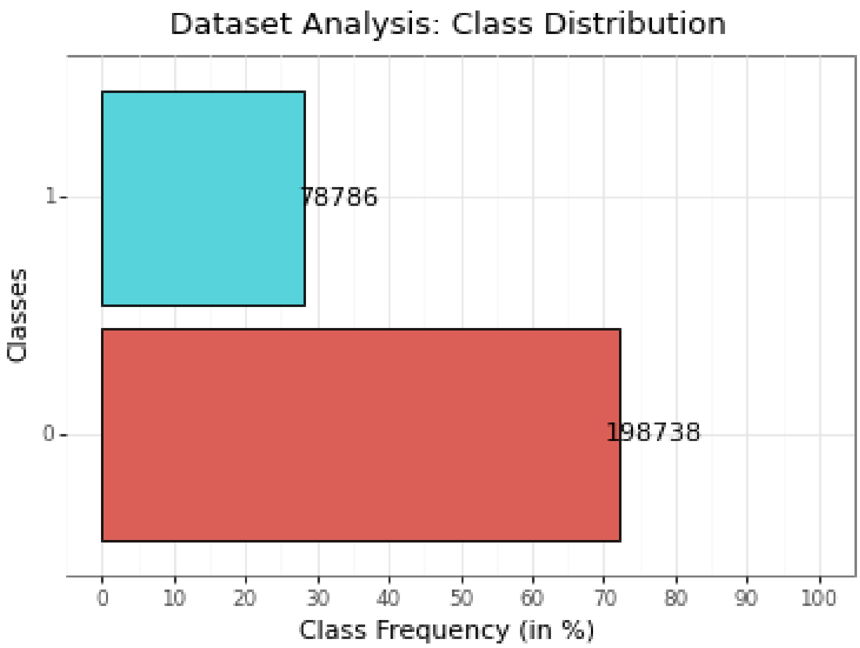



Figure 1: Class distribution: Breast Histopathology Images

## 2 Pre-processing

Images are originally associated to certain patients. This information is discarded during pre-processing and not respected for splitting. Instead, images are simply divided into classes 0 and 1. That may result in a patient imbalance among train, validation and test set.

## 3 Training

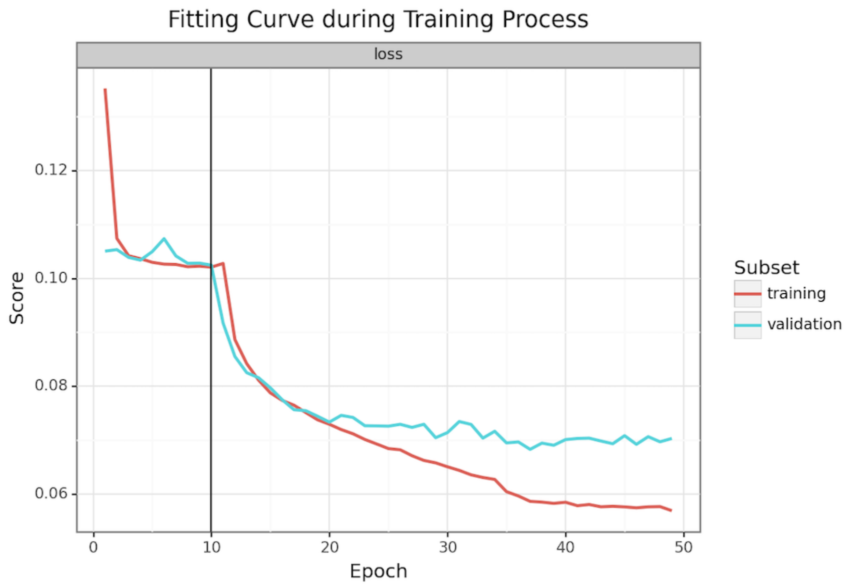



Figure 2: Training: Breast Histopathology Images

Usually, the unfreezing of the model's weights in episode ten results in peaking training and validation loss. That does not occur for this instance. Instead, opening up the model's weights causes a rapid, yet smooth reduction of both losses.

# 4 Results


Figure 3: Metric overview: Breast Histopathology Images

F1 scores show a great classification performance for both classes. Also, it is noteworthy, that the ROC curve (see figure 4) does not contain any visible steps due to the massive number of samples (see figure 1). The beautifully shaped ROC curves result in approximately 97 % area under the curve (AUC in figure 3), which confirms the model's reliable classification performance.

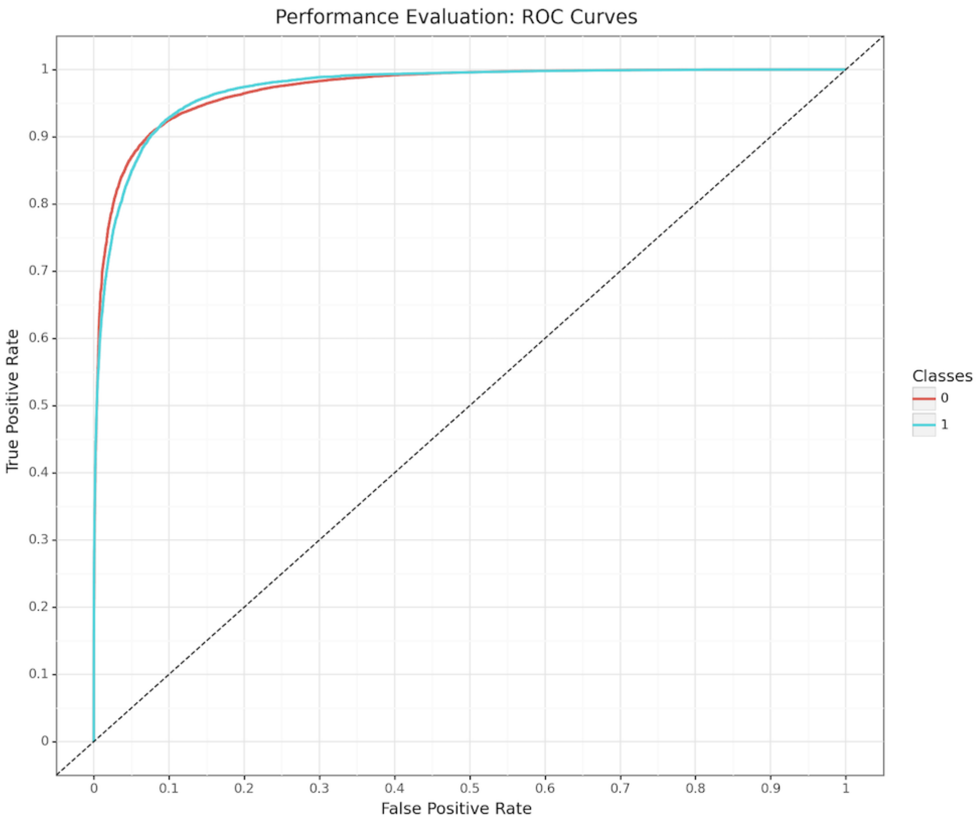



Figure 4: ROC curve: Breast Histopathology Images

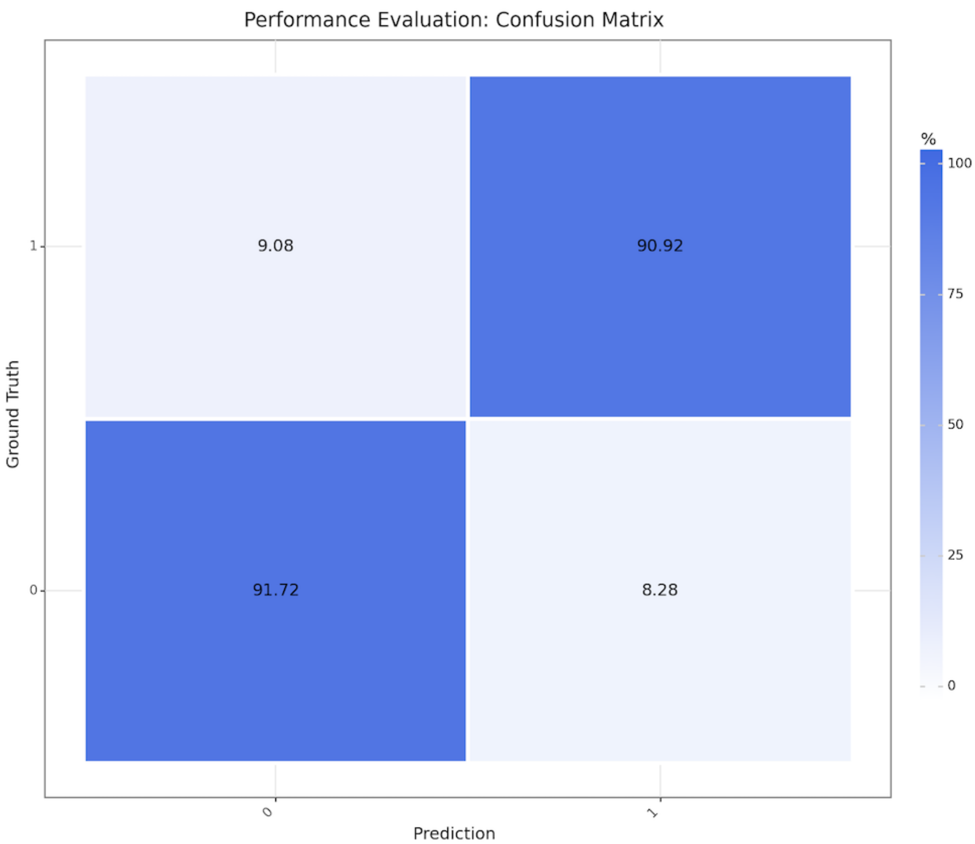



Figure 5: Confusion matrix: Breast Histopathology Images

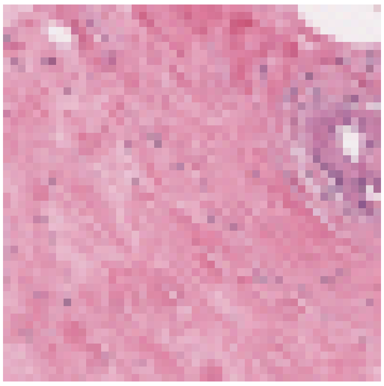
# 5 XAI

All images in this dataset are sized 50 by 50 pixels. Hence, Grad-CAM images are not ideal to understand the machine-made decision. We can see, that the classification of image **12751 idx5 x2301 y1801 class0** was based upon a small cluster of pixels in the top right corner. For the classification of image **10273 idx5 x1651 y1951 class1** more pixels seem to be relevant. Both were correctly classified with a solid confidence of over 80 %.

**Image:**
12751_idx5_x2301_y1801_class0

**Class:** 0

**Classified as:** 0 (83.9 %)



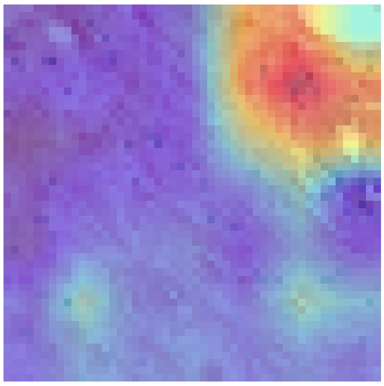
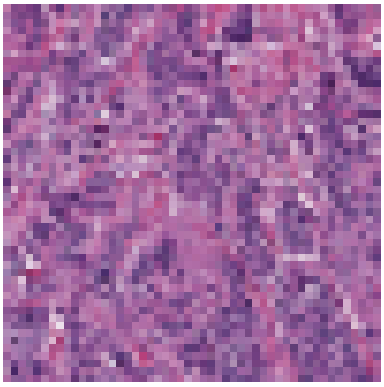
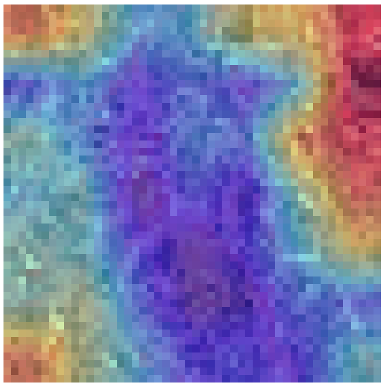


**Image:**
10273_idx5_x1651_y1951_class1

**Class:** 1

**Classified as:** 1 (83.6 %)





# 6 Summary

Although this dataset is very imbalanced (see figure 1), the model was able to achieve a great classification performance, probably due to the immense amount of samples. It appears, that not balancing the patients among classes did not negatively impact the model's performance. It would however be interesting to compare this primitive sampling with a more sophisticated, patient-balanced approach.

Grad-CAM images aim to highlight parts within an image. If the image itself is so small, that hardly any details can be recognized, the Grad-CAM mark-up cannot provide any insights.