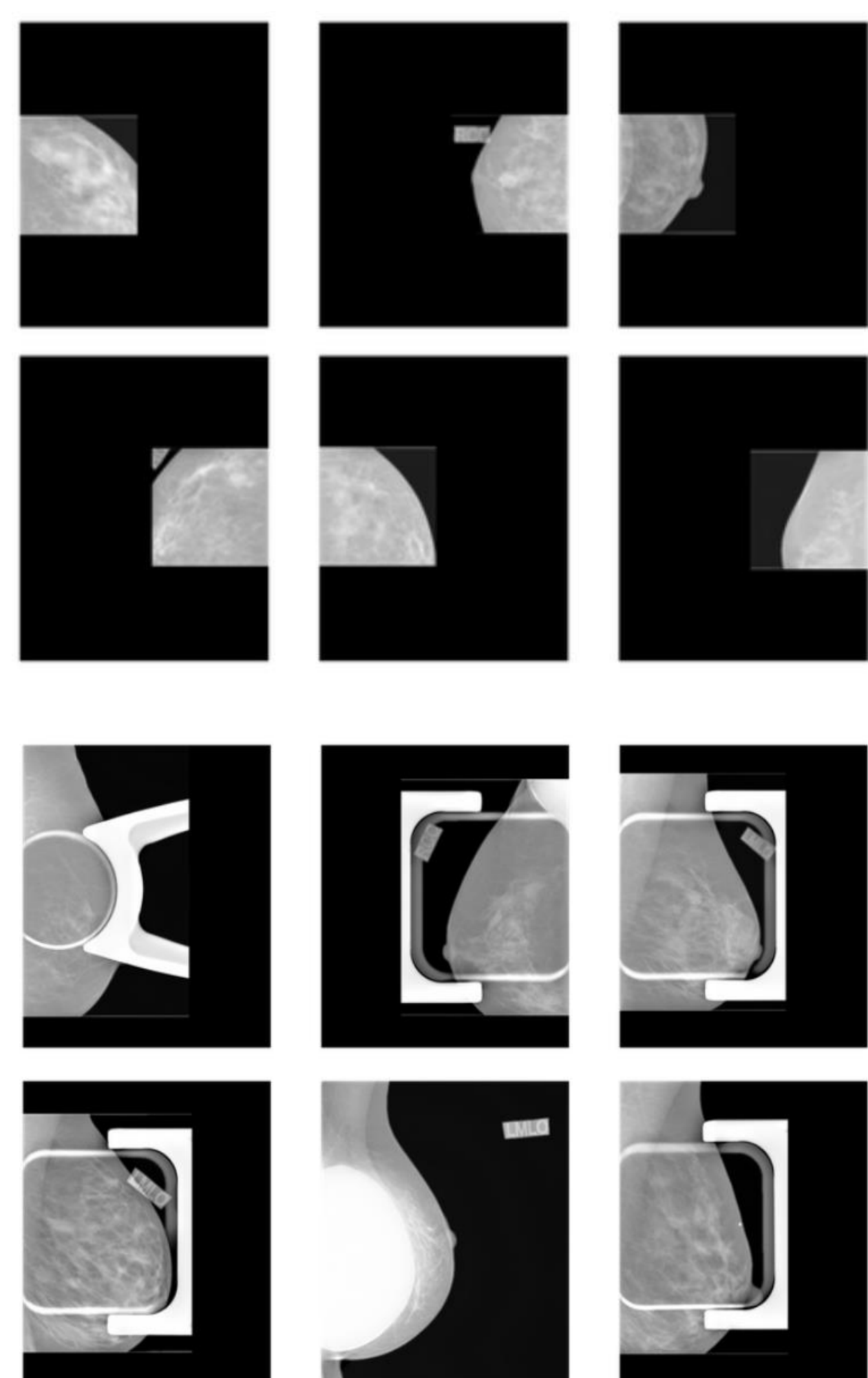


Outlier Detection for Mammograms

Ryan Zurrin, Neha Goyal, Pablo Bendiksen, Muskaan Manocha, Dan Simovici, Nurit Haspel, Marc Pomplun, Daniel Haehn

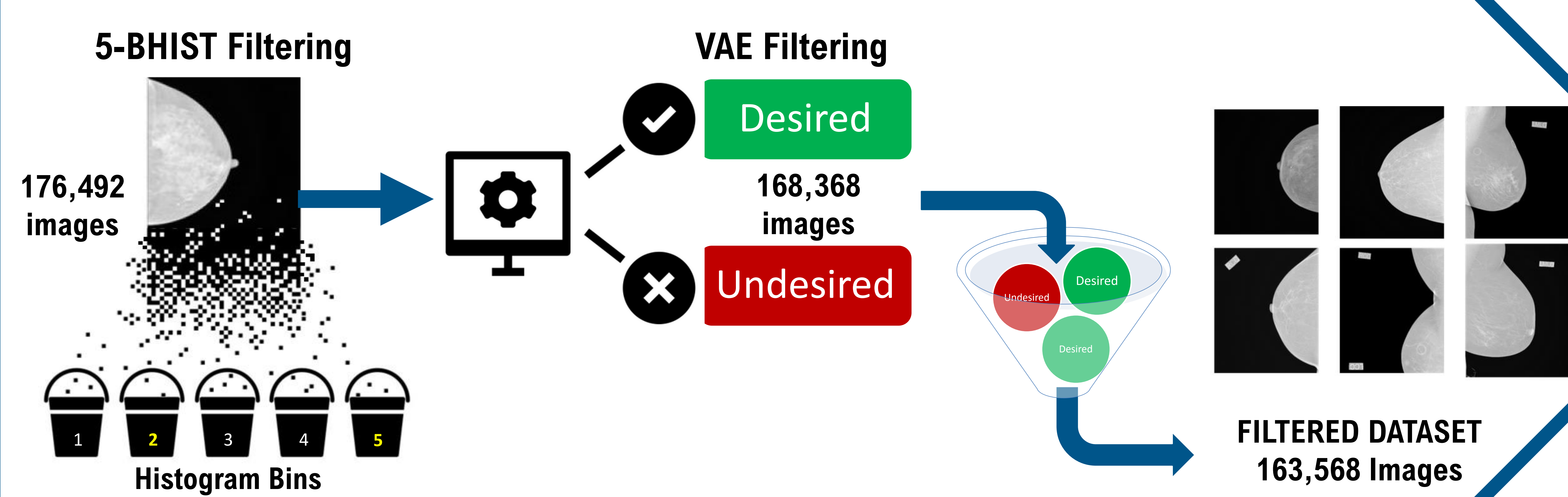
Example undesired images



Advanced Machine Learning tools depend on **GOOD Data**. Our Pipeline can improve the percentage of good data in mammography datasets.

- **Two-stage approach to remove undesirable images**
 - **Stage 1: 5-Bin Histogram Filtering (5-BHIST)**
 - **Stage 2: Variational Autoencoder (VAE)**

- 12,924 images removed
- 5.93x fewer unwanted images
- Improved to 1% from 5.5% of unwanted images.



- Method established after **rigorous testing** of 26 unsupervised outlier detection algorithms
- **Best overall results** with an average **F1 Score** of 0.8772, improved to **0.8862** with VAE



Check out the paper on OpenReview.net



<https://openreview.net/pdf?id=4E93Xdg98u>

Code available on GitHub



<https://github.com/mpsych/ODM>

