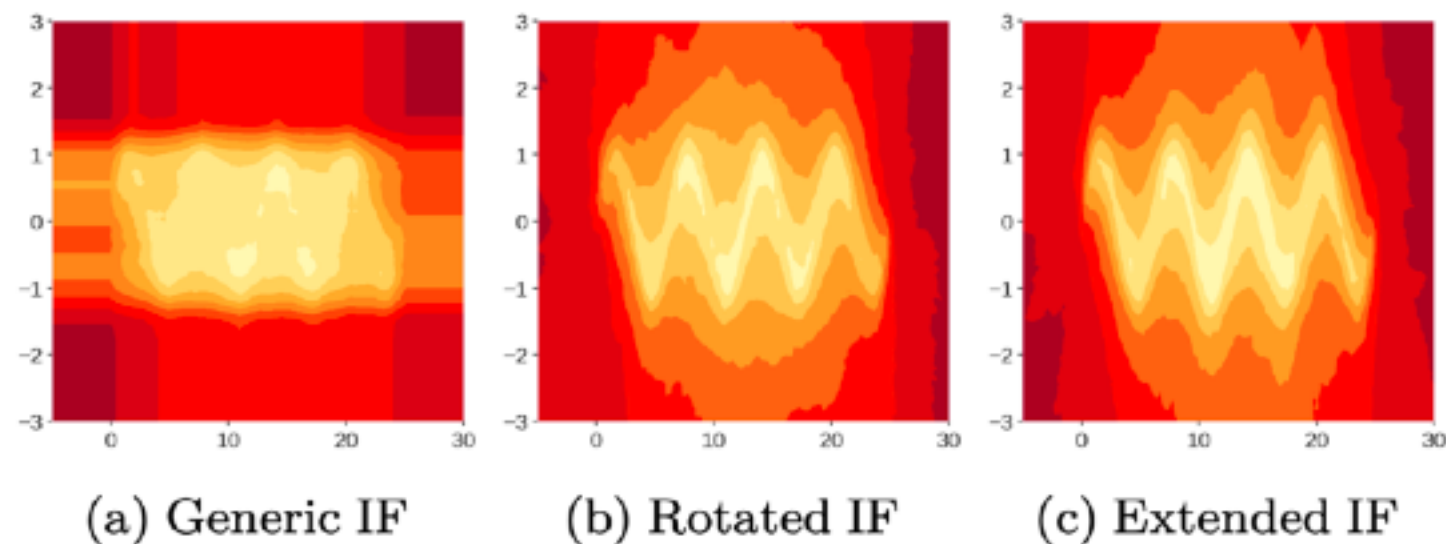


(Extended Isolation Forest)

# EIF to the rescue!

\* essentially it attempts to fix this inconsistent anomaly scoring region issue



# What am I doing?

## IF vs EIF

\* a battle for the history books

1. Thyroid dataset (2.5% known outliers)

2. How does IF perform?

3. How does EIF perform?

performance  
=  
# of outliers correctly identified

4. But what about those pesky scoring regions?!

5. Always a happy ending, r-right??

\* yes, the border between classification and outlier detection seems a bit murky here, but remember outlier detection is meant to be an unsupervised problem; it's just useful to know the true nature of outliers in this case to have some metric of performance