

# Attention Pattern Analysis in Flag Classification

Standard CLIP  
Accuracy: 18.0%



RS5M Fine-tuned (artifact: majority-class collapse)  
Accuracy: 72.6%



RS5M + Economic  
Consolidation  
Accuracy: 94.8%



Attention Intensity  
1.0  
0.8  
0.6  
0.4  
0.2  
0.0

- Focuses on building edges (65%)
- Weak flag symbol attention (15%)
- Background-biased features

- Strong flag region focus (75%)
- Symbol recognition improved
- Limited spatial context

- Hierarchical symbol + context
  - Economic consolidation
  - Optimal 169x improvement