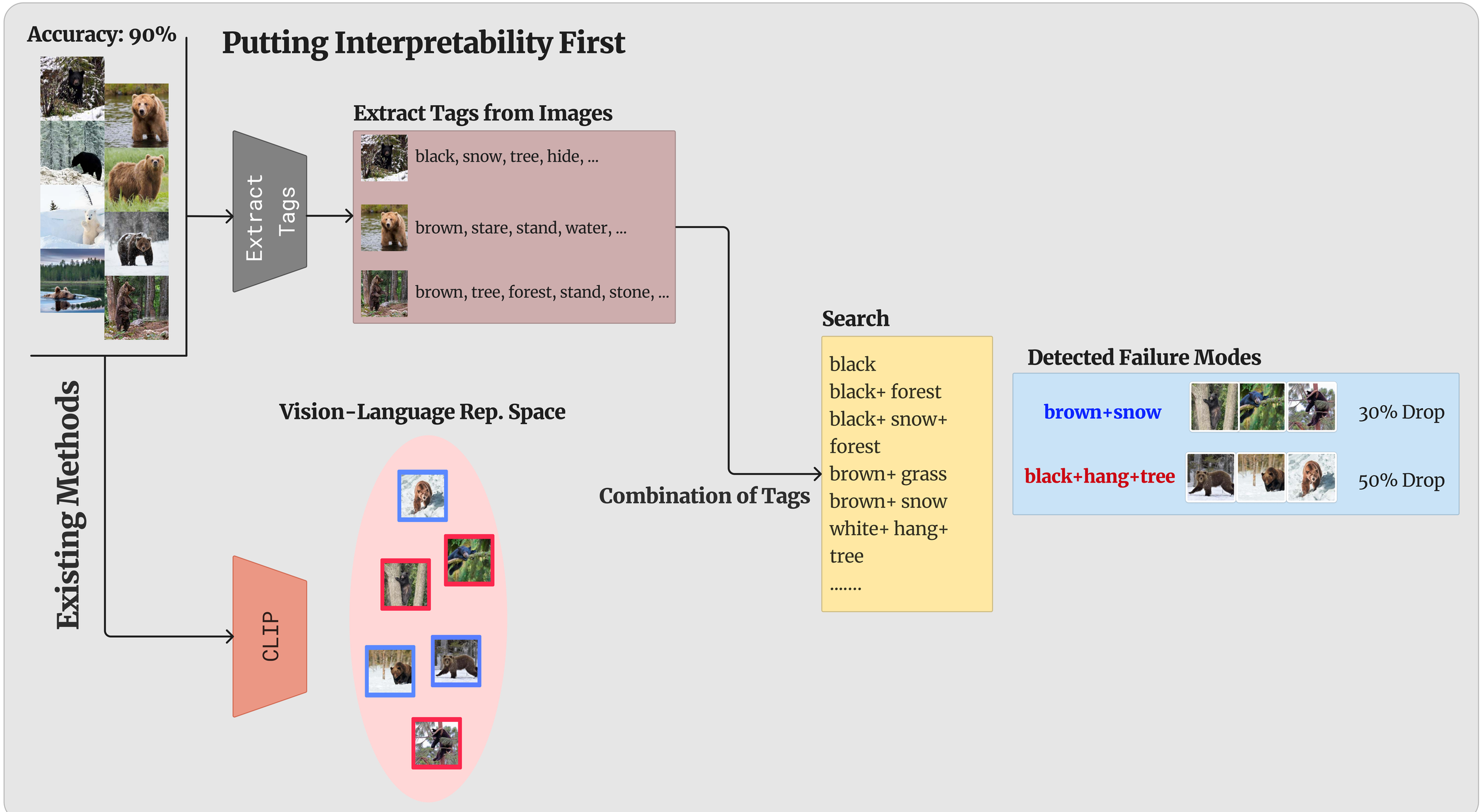
Accuracy: 90%
Putting Interpretability First
Extract Tags from Images
Extract Tags
black, snow, tree, hide, ...
brown, stare, stand, water, ...
brown, tree, forest, stand, stone, ...
Existing Methods
Vision-Language Rep. Space
CLIP
Combination of Tags
Search
black
black+ forest
black+ snow+ forest
brown+ grass
brown+ snow
white+ hang+ tree
.....
Detected Failure Modes
brown+snow
30% Drop
black+hang+tree
50% Drop