The article I am interpreting is titled:

"DynamicViT: Efficient Vision Transformers with Dynamic Token Sparsification"

This paper introduces a framework to improve the efficiency of vision transformers by dynamically pruning less informative tokens during inference, achieving significant computational savings with minimal accuracy loss.