

ABSTRACT

Due to subtle inter-class variations and high intra-class similarity, Fine-Grained Visual Classification (FGVC) presents significant challenges, often leading to attention bias and difficulty in learning truly discriminative features. Using counterfactual reasoning to measure the impact of attention, Counterfactual Attention Learning (CAL) was presented as a method to address this problem by comparing it with randomly generated fake attention maps. Different probability distributions help model these false attentions, allowing the network to learn more efficient and objective attention patterns. However, the natural randomness of these distributions can bring stochastic variability that limits model stability and performance. This study proposes the Annealed Counterfactual Attention (ACA) mechanism, which incorporates an annealing strategy into the CAL architecture, to address this issue. ACA enables the model to investigate several attention behaviors during the early training phase and progressively transitions towards actual attention, thereby improving its capacity to focus on pertinent and discriminative areas. This gradual improvement enhances both generalization capacity and attention quality. Two FGVC benchmark datasets, CUB-200-2011 and FGVC-Aircraft, were used for experiments. The proposed ACA model surpasses state-of-the-art methods, improving accuracy by 0.77% on FGVC-Aircraft and 0.09% on CUB-200-2011, while reducing inference time by 77.27 s and 11.65 s, without increasing parameters or Floating-Point Operations (FLOPs) compared to the baseline CAL model. This demonstrates the effectiveness of the annealing mechanism in stabilizing attention learning and enhancing classification accuracy.

Keywords: Annealing, Random Distribution, Counterfactual Reasoning, Attention Mechanism, Fine-grained Classification.