

Visual Recognition with Humans in the Loop

The authors claim that object recognition algorithms are not developed well enough to provide practical applications, albeit all the effort going to the object recognition task. Thus, the authors approach the problem from another angle and try to solve two basic questions to reason about what lies behind the lower performance of the objection recognition algorithms than expectations. The authors propose a “human-in-the-loop” method for object classification where the distinction line between two phenomena “what is easy for human to recognize” and “what is easy for a computer to recognize” is drawn. The authors aim at leveraging long-term memory of a computer system, while using humans as the supervisors as it learns how to distinguish classes finely-related to each other.

The paper formulates the problem by employing famous 20 questions games where the whole system is required to distinguish the object class by asking 1 questions for each step, and less than 20 questions in total. The method is essentially an application of Bayes Framework where whenever a new answer provided by the user it is regarded as a new observation and it is incorporated into prior distribution with Bayes’ formula. Along with user responses, the computer vision output is incorporated in similar manner. One of the significant property of the contribution is that any multi-class object recognition algorithm can be plugged into the system if it produces probabilistic output. The ultimate goal of the system is to reduce the number questions needed for any classes to be distinguished.

Results and Conclusions