

1 Abstract

For some people social networking websites are a large part of social life. Many of these people may not realise how much sensitive information they are sharing on these sites, and how easily identifiable they are from a starting point of as little as a picture of their face. This paper will examine the viability of unconstrained facial recognition in tandem with other open source information to identify social media targets. This capability mirrors the work done by the NSA and GCHQ, and should provide some insight in to how difficult (or otherwise) it is to identify a person given minimal information. The program developed is a command line tool which takes multiple inputs and will return the 100 closest matches to be sifted through manually. Since this paper has obvious ethical implications, test accounts will be generated to provide statistics on accuracy.

Unconstrained facial recognition has proven to be an exceptionally difficult problem in computing. While some success has been achieved using controlled samples with cooperative subjects, facial recognition in the context of social media has not been significantly explored in academia. Several confounding factors exist which make this problem much harder, such as variance in subjects' pose, ambient light and facial occlusion. To increase the accuracy of the tool developed, other identifying information has been used e.g. subject's name, home town etc.