C9 Human Versus Computer: Age Estimation Applied to Child Pornography

Mateus D.C. Polastro, MSc, Brazilian Federal Police, Campo Grande, Mato Grosso do Sul 79110-503, BRAZIL*

Learning Overview: After attending this presentation, attendees will understand some of the challenges related to the identification of child pornography images and will learn some promising technologies that can be used against the advancement of this type of crime.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by providing a comparative study of the ability to estimate people's ages using computer vision and pattern recognition versus the human ability to predict the age. In addition, this presentation will open a discussion concerning automatic age estimation for cases of child pornography and its impact.

Worldwide law enforcement is struggling to combat child pornography. With the increasing use of computers, cell phones, the internet, Deep Web, and social networks, among others, it is very common to find photos and videos stored in digital format on computers and cell phones. Thus, in investigations of child pornography, it is necessary to analyze thousands or even millions of images.

When an investigator is looking for child pornography images, the first step is to find images containing nudity or sex scenes. A person can perform this step visually or even in an automated manner.[1,2,3] After that, the task is to identify the participation of at least one child in the image. The presence of adolescents is also important in child pornography cases, since there are punishments in most countries. The possibility of automating the age estimation step using available age estimation algorithms may be an alternative in helping to decrease digital forensic backlogs related to child pornography. In addition, it can minimize the negative impact of exposure to investigators from child pornography images.

To identify the human ability to estimate age from photos, a survey was conducted in which respondents would estimate the age of 47 people displayed in pictures.[4] The photos were of people of both genders who were between 0 and 25 years of age. The images were in uncontrolled scenes, contained no nudity, and the respondents would estimate the age based only on face and body proportions.

With the goal of comparing the automatic age estimation with the results obtained in the study described above, the same 47 images were submitted to some state-of-the-art age estimation services: Microsoft's® how-old.net, Face ++™, and VeriLook SDK™.[5-7] The result obtained by each algorithm was then compared to the answers provided by the survey respondents. It is important to note that in cases related to child pornography, generally, the main goal is to locate the presence of a child or adolescent in the image, and it is not necessary to define the exact age. Thus, three comparison groups were created: (1) people under 14 years of age; (2) people between 14 and 18 years of age; and (3) people over 18 years of age. The results of the estimates made by the people and those made by the algorithms were considered an error when the person's real age was in one of these groups but was not estimated in the correct group. Additionally, in some cases, the algorithms did not detect faces in any images, so it was not possible to estimate the age, and these were not counted as errors.

The findings of this research demonstrate that age estimates made by the survey respondents were more accurate than those performed by the tested algorithms. On the other hand, the algorithms did not classify any person of group 3, in group 1, or in group 2; that is, no false positive, unlike the estimates made by the survey respondents, who misjudged 33% of the cases. The highest error rates occurred in group 2, both by algorithms and by the survey respondents. It was expected the error rate would be higher in this group since adolescents may often have characteristics that confuse them with adults.

The results of this study suggest that algorithms have not yet outperformed humans in estimating people's age. Nevertheless, there has been much progress in the accuracy of the classifiers, and it is already possible to use them in real cases in which false negatives are accepted. Detailed results will be shared at this presentation and recommendations for further research will be discussed.

Reference(s):

1. Polastro, M.D.C.M. & Eleuterio, P.M.D.S. 2010. *NuDetective: A Forensic Tool to Help Combat Child Pornography Through Automatic Nudity Detection.* 2010 Workshops on Database and Expert Systems Applications, pp.349–353.
2. MediaDetective. <http://www.mediadetective.com>.
3. Snitch Plus. < http://www.hyperdynesoftware.com>.
4. Murphy C.A. (2012) The Role of Perception in Age Estimation. In: Gladyshev P., Rogers M.K. *Digital Forensics and Cyber Crime*. ICDF2C 2011. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 88. Springer, Berlin, Heidelberg.
5. Microsoft®. <https://how-old.net>.
6. Face++™ api. <http://www.faceplusplus.com>.
7. VeriLook SDK™. <http://www.neurotechnology.com/verilook.html>.

Age Estimation, Child Pornography, Nudity Detection