

ABSTRACT

CAPTCHA is a test program that automatically test and distinguish between human and computer with the goal to set up access to the website. CAPTCHA is shown in text, images or audio. They are difficult to be recognized by the computer but can be recognized by humans. CAPTCHA is used to prevent spam that tried to enter the system. But over the times, people worry if there is a program that can solve CAPTCHA challenge. Therefore, the author conducted a text-based CAPTCHA solving challenges with Gimpy types that have been implemented in the website www.hotmail.com to check how robust CAPTCHA defenses against a website .

The security system works by displaying Gimpy CAPTCHA text type which has in the form of upper / lower case and numbers with warping for distortion. In the breaking scheme, the initial step includes preprocessing pursued by converting color to black and white, eliminate small pixels and then cropping. Followed by the character segmentation process to separate the pixels based on their distribution which mutually exclusive of each other, then the process of normalization include rotation, resize the character size and morphological operations. The final step is feature extraction process using square method and classification using Mean Square Error (MSE) .

From this simulation produced the best accuracy for segmentation is 79,167% and CAPTCHA challenges are solved as many as 15 samples from a total of 120 sample trials resulting in an accuracy of 12,5 % CAPTCHA file with character accuracy of 70.64 % . The average computation time for each CAPTCHA recognition process lasts for 0.156 seconds .

Keywords : *CAPTCHA, Artificial Intelligence, spam, preprocessing, website, cropping, resize, square, Mean Square Error*