|  |
| --- |
| **착시현상을 이용한 이미지 CAPTCHA 프레임워크 제시** |
| 한준희1  **1민족사관고등학교** |
| **Designing a novel image CAPTCHA framework based on illusion** |
| **Jooney Han1**  **1Korean Minjok Leadership Academy, Hoengseong-gun, Gangwon-do, Korea** |
| **요 약** 최근 급속도로 발전하는 Vision AI 및 Multimodal AI를 악용하여 CAPTCHA의 파훼법이 개발됨에 따라 심각한 보안 문제가 발생하고 있다. CAPTCHA는 웹사이트에 접속하는 자동화 프로그램을 차단하기 위해 개발된 테스트로, 온라인 투표, 회원 가입 등 사람만이 접근할 수 있어야 하는 기능을 보호하기 위해 사용된다. 하지만 최근 발전하는 AI 기술을 사용하여 이를 자동으로 파훼하고, 크롤링 프로그램을 통해 데이터를 무단으로 수집하거나 악성 데이터를 주입하는 등의 피해가 발생하고 있다. 이에 본 연구에서는 이미지의 착시 현상을 활용하여 AI는 파훼 하지 못하지만, 사람은 문제없이 해결할 수 있는 새로운 CAPTCHA 프레임워크를 제시한다. 상술한 AI들은 이미지 내에 있는 사물들 뿐만 아니라 전체적인 분위기와 구도까지 파악하고 분석 할 만큼 성능이 발전된 상태이다. 따라서 위의 문제를 해결하기 위해 기존의 방식과는 다른, 착시 현상을 이용한 CAPTCHA 프레임워크를 개발하고, 이를 각각 AI와 사람이 어떻게 해결하는지 비교하였다. 그 결과 AI는 테스트를 파훼 하지 못하며, 사람은 큰 문제없이 통과함을 확인했다. 이에, 본 연구에서 제시하는 CAPTCHA 프레임워크가 AI 및 자동화 프로그램의 웹사이트 무단 접근을 효과적으로 방지함을 확인한다. |
| **Abstract** Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper Guidelines for KSEF research paper  **(서체:신명조, 크기:8.5pt, 양쪽정렬, 8줄 또는 200단어 이내)**  **Key Words: CAPTCHA, Illusion, Multimodal AI, Stable Diffusion ControlNet** |

1. 서론

1.1 연구 배경

CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart)는 HIP(Human Interaction Proof) 기술 중 하나로, 웹상에서 데이터를 무단으로 수집하거나, 사람만 접근할 수 있어야 하는 페이지를 접근하여 악성 데이터를 주입하는 등 다양한 피해를 끼치는 자동화 프로그램을 차단하고자 개발되었다. 캡차 (CAPTCHA)는 질문에 맞는 이미지 선택, 논리 문제 해결, 왜곡된 글자 인식 등 컴퓨터는 해결하지 못하지만, 인간은 매우 쉽게 해결할 수 있는 문제를 제시하여 인간을 구분한다. 하지만, 최근 급속도로 발전하고 있는 AI를 악용하여 높은 정확도로 이를 파훼하는 solver 프로그램이 등장하기 시작했다. 왜곡된 텍스트를 사용하는 캡차는 CNN 모델에 의해 파훼된지 오래이며, 비교적 해결하기 어려운 이미지 선택 캡차도 최근 객체 인식 AI에 의해 파훼된 상태이다. [1] 또한, Flamingo, CLIP, VisualBERT, GPT-4, PaLI 등 텍스트, 이미지, 오디오 등 다양한 입력을 동시에 처리할 수 있는 Multimodal AI의 등장으로 AI가 대부분의 캡차의 질문을 파악하고 해답을 도출할 수 있게 되었다. [2-6] 따라서 본 연구에서는 원래 의도와는 다르게 웹상에서 자동화 봇을 차단하지 못하고, 사용자들의 접근성만 저하시키는 현존하는 캡차의 문제점을 해결하고자 하였다. 이를 위해 착시 현상을 사용하여 AI는 해결하지 못하지만, 사람은 쉽게 해결할 수 있는 새로운 캡차 프레임워크를 개발하게 되었다.

1.2 선행 연구

캡차는 1997년 카네기 멜런 대학교에서 최초로 연구되어, 현재 웹상에서 보편적으로 사용되고 있는 캡차 서비스들로 발전되었다. 하지만 인터넷 상에서 사용되는 캡차의 96.48%를 차지하고 있는 서비스인 reCAPTCHA도 앞서 언급했듯이 파훼 프로그램이 존재하며, 이외에도 hCaptcha, Cloudflare Turnstile 등 다양한 캡차 프로그램의 파훼법이 등장한 상태이다. [1, 7-8]

이렇듯 캡차의 파훼법이 점차 개발됨에 따라, 이를 방지하기 위한 논의도 활발히 진행되었다.

hCaptcha, reCAPTCHA, Cloudflare Turnstile 등으로 발전되었다.

2. 이론적 배경

2.1 연구 배경

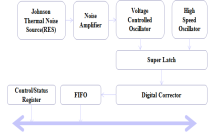
2. 본론**(서체:중고딕, 크기:11pt)**

2.1 각 제목**(서체:중고딕, 크기: 9.5pt)**

2.1.1 각 제목**(서체:중고딕, 크기: 9pt)**

여기부터 논문을 작성한다.

Fig.와 Table의 작성은 Fig. 1과 같이 작성 하며, 모든 내용과 제목은 국문 및 영문으로 작성한다.



[Fig. 1] Figure title

**(아래캡션, 서체:신명조, 크기: 8pt, 양쪽정렬,)**

다음과 같은 현상은 [Fig. 1]에서 보는 바와 같은 순서도에 따라 설명할 수 있는데...

|  |  |
| --- | --- |
| EMB000227c805ab  (a) | EMB000227c805ac  (b) |

[Fig. 2] Figure title

(a)Figure title (b)Figure title

**(아래캡션, 서체:신명조, 크기: 8pt, 양쪽정렬,)**

구조물에 대해서는 [Fig. 2]와 같은 수치해서 통해 증명할 수 있다. 수치해석에 사용한 물질은 [Table 1]에서와 같은 물질 특성을 가지고 있다.

[Table 1] Table title**(위 캡션, 서체:신명조, 크기: 8pt, 양쪽정렬, \*표 내용 : 8pt, 중고딕,)**

|  |  |  |  |
| --- | --- | --- | --- |
| Spec. | | Respondents | Percentage |
| Gender | Male |  |  |
| Female |  |  |
| Recognizing multicultural families | Yes |  |  |
| Maybe |  |  |
| No |  |  |
| All | |  |  |

3. 결론

여기부터 논문을 작성한다.

References

[1] Hossen, Imran et al. “An Object Detection based Solver for Google's Image reCAPTCHA v2.” International Symposium on Recent Advances in Intrusion Detection (2021).

[2] Alayrac, Jean-Baptiste, et al. “Flamingo: A Visual Language Model for Few-Shot Learning.” 2022.

[3] Radford, Alec, et al. “Learning Transferable Visual Models From Natural Language Supervision.” 2021.

[4] Li, Liunian Harold, et al. “VisualBERT: A Simple and Performant Baseline for Vision and Language.” 2019.

[5] OpenAI. “GPT-4 Technical Report.” 2023.

[6] Chen, Xi, et al. “PaLI: A Jointly-Scaled Multilingual Language-Image Model.” 2023.

[7] Hossen, Md Imran, and Xiali Hei. “A Low-Cost Attack against the HCaptcha System.” 2021 IEEE Security and Privacy Workshops (SPW), IEEE, 2021, doi:10.1109/spw53761.2021.00061.

[8] CAPTCHA Usage Distribution on the Entire Internet, <https://trends.builtwith.com/widgets/captcha/traffic/Entire-Internet> 2023.12.15

|  |  |
| --- | --- |
| 한 준 희 (Jooney Han) | |
| 의류, 사람, 인간의 얼굴, 턱이(가) 표시된 사진  자동 생성된 설명 | • 2022년 2월 : 가원중학교 졸업  • 2022년 3월~현재 : 민족사관고등학교 재학 |
| <관심분야>  Computer Vision, Artificial Intelligence, Information Security | |