머신러닝 기반의 온라인 컴퓨터 게임 핵 검출

이세훈*, 우찬혁*, 김기태*, 정석주O, 박전진**
O인하공업전문대학 컴퓨터시스템과,
*인하공업전문대학 컴퓨터시스템과,
**㈜웰비아닷컴
e-mail: seihoon@inhatc.ac.kr*, dncksgur123@gmail.com*, gth3641@naver.com*,
gotter9179@gmail.comO, apuro@wellbia.com*

Machine Learning based Online Computer Game Hack Detection

Se-Hoon Lee*, Chan-heok Woo*, Gi-Tae Kim*, Seok-Ju JeongO, Jun-Jin Park**
ODept. of Computer Systems, Inha Technical College,
*Dept. of Computer Systems, Inha Technical College,
**Wellbia, Ltd. Co

● 요 약 ●

본 논문에서는 현재 운영되고 있는 온라인 게임에서 실력을 겨루는 형태의 경쟁적인 온라인 게임들에서 사용되어지고 있는 게임 핵이 게임에 미치는 영향을 제시한다. 그리고 게임 핵을 검출하기 위한 객체 인식 기술로 실시간 정보 획득이 가능한 YOLOv3 알고리즘을 사용하였다. 이는 속도가 빠른 객체인식 기술이며 이미지 속 물체의 외관 뿐만 아니라 전체적인 컨텍스트까지 학습을 진행한다. 그리고 나아가 게임 핵 검출을 위한 개발 및 운영적 측면에서 어떻게 지원돼야 하는 등의 내용을 제시한다.

키워드: 게임핵(Game hack), 객체인식(Objection Recognition), 욜로(YOLO)

I. Introduction

컴퓨터 온라인 게임에서 핵의 문제는 사회 문제가 될 만큼 큰 영향을 미치고 있으며 경제적인 문제까지 연결이 되어서 상당한 사회적 파장을 낳고 있다.

본 연구에서는 객체인식 머신러닝을 통한 컴퓨터 부정 프로그램을 자동으로 하는 시스템을 개발한다. 부정 프로그램이라 하면 흔히 핵(Hack)이라고 말하는 비인가 프로그램이다.

핵 검출을 위해 근래 많은 발전이 되어 온 머신러닝의 객체 인식 기술을 이용한다. 객체 인식 기술의 활동 범위는 로봇의 비전, 가전제품, 이미지 검색, 자율주행 등 새로운 분야와 실생활의 분야까지 다양한 분야에서 중요한 기술로서 사용되고 있다.

본 논문에서는 기존의 Yolo를 이용해 전이학습을 통해서 객체 인식을 정확도를 높이고 실시간으로 검출 핵을 통보할 수 있는 시스템을 개발한다.

II. The Proposed Scheme

1. Project

상대팀에 핵 유저가 있는 경우는 핵을 이용한 강력한 가능에 직접적인 피해를 입는 경우이며, 같은 팀에 핵 유저가 있는 경우 또한 게임의 재미를 망가트리는 경우로 피해 사례라고 볼 수 있다.

라이엇 게임즈에서 운영하고 있는 <리그 오브 레전드>에서 활동했었던 프로게이머 'Bokgu' 정재우 선수가 핵을 사용해서 게임을 하다가 적발되어 계정 사용을 영구히 정지 당하고 일반 유저들뿐만이 아닌 프로게이머 조차 핵을 사용하여 제재를 받는 등 무분별한 핵 사용이 이어지자 기존 사용자들은 핵에 점차 무관심해지며 게임 내에서 핵 사용자를 만났을 경우 신고를 해도 소용없다고 판단하여 신고를 하지 않거나 게임을 더 이상 플레이 하지 않게 되는 경우도 발생한다.[3]

게임사에서는 지속적으로 게임의 업데이트 내용과 함께 공지사항을 통해 불량 이용자 차단이 어떤 식으로 이루어지고 있는지와 함께 어떤 이용자들이 총 몇 명 차단되었는지에 대한 내용을 공개해 유저들에게 신뢰를 주기 위해서 노력하고 있다.

해당 프로젝트에서는 FPS(1인칭 슈팅게임 : First-Person Shooter) 기반의 게임중 배틀로얄 장르의 배틀 그라운드(이하 PUBG)를 택했

다. 해당 게임을 택한 이유는 FPS장르의 게임이 특히나 핵 사용자의 수가 몰려있고 또한 취약하기 때문이다. 그 중 세계적으로 많은 유저를 보유한 PUBG를 선택하였다.

2. Making Model

YOLO v3을 통해 모델을 생성하는 과정은 크게 데이터 셋 정규화, 예측으로 나눌 수 있다.

2.1 데이터 셋 정규화

YOLO에서 데이터 셋을 정규화 시키는 과정은 이미지를 라벨링 시켜주는 작업이다.

[Fig. 4.]는 프로젝트에서 진행한 데이터를 라벨링 하는 과정이다. 여기에서 주의해야 할 점은 데이터를 라벨링 할 때 노이즈를 일으킬 만 한 데이터는 포함시키면 좋은 결과값을 받을 수 없다는 것 과 충분한 양의 데이터를 라벨링 하는 것이 중요하다.

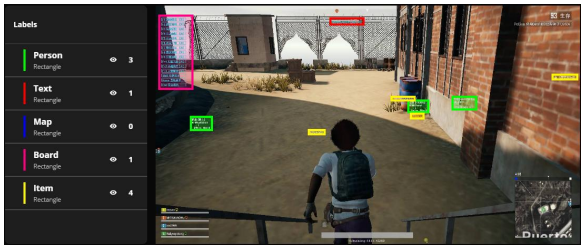

Fig. 4. Data Labelling

또한 데이터 셋 정규화와 관련해서 아래와 같은 사항들을 유의해야 한다.

대표성이 없는 훈련 데이터는 한 모델을 훈련시키는데 이러한 데이터가 들어가게 된다면 샘플링 잡음(sampling noise)가 생기고 큰 샘플 또한 표본 추출 방법이 잘못되었다면 샘플링 편향(smapling bias)이 일어날 수 있다. 이외에 데이터셋을 훈련함에 따라 너무 훈련량이 적으면 과소적합(under fitting) 으로 제대로된 예측을 수행할 수 없고 훈련량이 너무 많다면 반대로 과적합(over fitting)으로 훈련 데이터에는 너무 잘 맞지만 일반성이 떨어지는 모델을 얻게 될 수 있다.[6]

일반적으로 classes 하나당 2000회정도의 훈련을 추천하고 있으며 Loss 1%이하 정도면 사용할 만한 수준의 가중치를 만들 수 있다

2.2 예측

만든 가중치 파일로 데이터를 예측 할 때 만든 목적의 기준을 제대로 세워야 한다. 이 프로젝트를 예로 들어보자.

만든 가중치 파일을 테스트 할 때 생각하고 있던 모든 클래스를 완벽하게 찾지만 해당 클래스가 없는 데이터를 주었을 때 가끔씩 존재하지 않는 클래스를 찾아버리는 가중치 파일 1 이 존재하고

데이터를 주었을 때 생각하고 있던 모든 클래스를 완벽하게 찾지는 못하지만 해당 클래스가 없는 데이터에서는 아무 클래스도 못찾는 가중치 파일 2가 존재한다고 할 때 사용 목적에 따라 둘 중 하나를 잘 선택해야 한다. 사람이 다가오면 사람을 인식해서 작동해야하는 비교적 낮은 보안수준의 자동 문 같은 경우 사람을 인식하는 것이 중요하기 때문에 가중치파일 1을 선택하는 것이 이롭지만 해당 프로젝트에서는 부정 프로그램을 사용하는 플레이어만을 측정하는 것이 중요하기 때문에 가중치파일 2와 같은 가중치 파일을 선택하는 것이 이롭다.

III. Conclusions

온라인게임 시장의 가치가 떨어지지 않는 이상, 온라 인게임핵은 악성코드 제작자에게 황금알을 낳는 거위 가 될 것이고 계속해서 생산될 것이다. 게다가 루트킷, 부트킷 등의 고급 기술 등이 포함되고 있으며, 앞으로 도 기술적으로 많은 진보를 이룰 것으로 보인다.

REFERENCES

[1] Deep Learning for Generic Object Detection: A Survey 1809.02165v4 p.3
[2] A traffic light recognition and color detection based on YOLOv3
[3] A Study on the Actual Use of Game Hack in Online Games
[4] Game hack in Online game and strategy
[5] You Only Look Once: Unified, Real-Time Object Detection