**〈 25-1 프로젝트 계획서 〉**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 교과목 명 | 예시) 캡스톤디자인 | | | |
| 학수번호 | 예시) CSE406-31 | | | |
| 프로젝트 명 | oooooooo | | | |
| 지도교수 | 이름 | | 소속 | 이메일 |
| ooo | | ooo | ooo |
| 팀원 (멘티) | 이름 | | 학과 (다전공) | 이메일 |
| 팀장 | ooo | ooo(ooo) | ooo |
| 팀원 | ooo | ooo | ooo |
| 팀원 | ooo | ooo | ooo |
| 팀원 | ooo | ooo | ooo |
| 멘토 | 이름 | | 소속 | 이메일 |
| ooo | | ooo | ooo |

\*다전공 학생은 필히 표시

프로젝트 명: ooooo

# 과제 개요

## 과제 선정 배경 및 필요성`

예시) 기존 경찰청 유실물 종합관리시스템은 습득한 분실물의 정보를 관리자가 매번 수기로 등록해야 하며, 이에 따라 글이 작성되기 전까지 분실자는 물건을 찾아 헤매야 하는 불편함이 있었음. 따라서 분실물 이미지로부터 분실물의 종류를 구별하는 모델을 도입하여 분실물 관리 시스템을 더 효율적으로 개선하려고 한다.

최근 AI 및 클라우드 산업의 발달에 따라, 데이터 센터에서 GPU와 같은 자원의 수요가 급격하게 높아지고 있다. GPU 자원의 효율적인 사용 및 예측 가능한 운용은 데이터 센터에 있어서 필요하다.

리소스가 수요에 비해서 과도하게 많이 할당이 된다면 그 리소스는 낭비가 되는 것이고, 반대로 리소스가 수요에 비해서 부족하게 할당이 된다면 막심한 손해가 발생할 수 있다.

이를 해결하기 위해서 시간에 따른 자원 사용량의 패턴을 학습하고 미래를 예측할 수 있는 시계열 데이터 기반의 예측 모델이 필요하다. 그리고 예측 모델을 도입하여 리소스 관리를 더 효율적으로 할당할 수 있도록 한다.

## 과제 내용 요약

예시) 경찰청에서 자주 분실되는 분실물 카테고리에 최적화된 정확도가 높은 YOLO 모델을 개발한다. 분실물 카테고리는 지갑, 카드지갑, 목도리, 안경, 장갑, 휴대폰, 카드, 무선이어폰, 우산 총 9가지이다.

Unlabeled data가 많을 때에는 데이터를 레이블링하는 것이 현실적으로 불가능하기 때문에 supervised learning을 사용하기 어렵다. 이때 semi-supervised learning 기법을 사용할 수 있다. 이 프로젝트에서 활용하는 경찰청 유실물 종합시스템의 이미지 데이터들의 annotation이 되어 있지 않으므로, semi-supervised learning 기법을 사용하는 것이 적절하다고 판단했다. Semi supervised object detection은 unlabeled data와 labeled data(bounding box annotation)을 이용해 학습을 진행하는 object detection 방식이다.

Semi-supervised learning의 가정들에 기반해 나온 여러 method들이 있다. 이 method들을 임의로 조합해 직접 구현하는 방법과, 연구를 통해 도출된 유명한 semi-supervised model에 대해 학습한 뒤 이를 이용하는 방법이 있다. 연구를 통해 도출된 유명한 모델들에는 self-training, multi-view training, Π-model, Temporal Ensemble, Mean Teacher, VAT, UDA, consistent teacher, efficient teacher 등이 있다.

우리는 unlabeled data를 pseudo-labeling하고, pseudo-labeled된 데이터셋을 YOLO 모델에 반복적으로 학습시켜 정확도를 향상시키는 방법(pseudo-label, meta pseudo-label, noisy student, co-training 등)을 활용해 semi-supervised learning을 진행하려고 한다. 시간의 여유가 있다면 다른 모델들과의 성능을 비교하여 최적의 모델을 선정하고 그 이유를 알아본다.

이 프로젝트를 통해서 GPU 자원 사용 데이터를 기반으로, LSTM 모델을 활용하여 시계열 예측 모델을 구축하고 그 모델을 기반으로 데이터 센터의 리소스 할당을 보다 효율적으로 할 수 있게 한다.

# 과제 목표 및 기대효과

## 최종 결과물 목표

### 정량 목표 (측정 가능한 것)

* + - * 예) 효율성 및 보안성의 20% 이상 개선

### 정성 목표 (측정 불가능 한 것)

* + - * 예) 웹 보안에 대한 이해도를 높이고, 웹 기반 시스템의 취약점을 분석하여 이를 보완
      * 특히, SQL Injection 공격 기술 및 보안 취약점을 분석하고, 새로운 보안 메커니즘과 방법을 개발하여 웹 시스템의 보안 강화에 기여

### 논문

* + - * 국내 학술 대회 논문 1편 제출

## 최종 결과물의 세부 내용 및 구성

### 예시) SQL Injection 공격 분석

* + - * Union-based, Boolean-based, Error-based 등 다양한 SQL Injection 공격 기법을 분석하고, 각기법의 특징과 작동 원리를 파악하여 이에 대한 대응 방안을 탐색한다.
      * 실제 사례를 조사하여 SQL Injection 공격의 실제 피해와 공격자의 공격 전략을 분석한다.
      * SQL Injection 공격을 시뮬레이션하고 발생하는 취약점에 대해 집중 분석한다.
      * 예시) 프로젝트의 최종 결과물은 크게 두 가지 주요 단계로 구성된다. 수면 뇌파 및 기타 생체신호 데이터의 분석 단계와 이미지 생성 단계이다.
      * 첫 번째 단계에서는 데이터 전처리 기법을 적용하여 Sleep EEG, ECG, EMG, EOG 등의 데이터를 전처리하고, 각 채널의 중요한 정보를 추출한다. OOOO
      * 두 번째 단계인 이미지 생성 과정에서는 앞서 분석 및 분류 모델을 통해 얻은 feature들을 바탕으로, 생성형 AI 기술을 활용해 사용자의 수면 상태를 상징하는 이미지를 제작한다.

### 연구 및 실험

* + - * 예시) 취약점 탐지 및 보안 강화를 위한 새로운 알고리즘을 개발하고, 웹 애플리케이션에 효과적으로 적용할 수 있는 기술적인 솔루션을 설계한다.
      * 개발한 메커니즘을 다양한 웹 애플리케이션과 공격 시나리오에 대해 실험하여 성능을 평가하고, 정량적 지표를 측정한다.

### 비교 및 분석

* + - * 새로운 알고리즘을 기존의 SQL Injection 대응 방법과 비교하여 공격 탐지율, 성능 및 효과적인 보안 강화 정도 등을 분석하여 차이점을 파악한다.
      * 분석을 토대로 새로운 알고리즘의 개선점을 도출한다.

## 기대 효과 및 활용 방안

예시) 현재 텍스트 발화를 듣거나 비디오를 시청하는 중에 측정한 뇌파를 분석하여 해당 텍스트가 설명하는 이미지 또는 시청한 비디오를 재현하는 기술은 개발되고 있지만 수면 상태에서 측정된 뇌파를 이미지화하는 방안에 관해서는 연구된 바가 거의 없다. 본 과제에서는 수면 뇌파 데이터를 이미지화하는 창의적인 아이디어를 도출하고, 실제로 이미지를 생성하고자 한다. 이때 주로 활용할 데이터는 뇌파 데이터인 EEG이지만, EMG, EOG, ECG 등 다른 생체 신호도 함께 사용함으로써 보다 다양한 공간 정보를 활용하여 풍부한 요소를 담고 있는 이미지를 생성할 수 있을 것이다. 이를 통해 수면 뇌파를 어떻게 이미지화할 것인지에 대한 다양한 시각을 제시할 수 있고, 뇌파 시각화에 관한 보다 발전된 논의와 연구를 활발히 끌어낼 수 있을 것이다. oooo

시각화된 이미지는 다양한 방법으로 display할 수 있는데, 이것을 여러 가지 재밌는 방법으로 응용한다면 예술적으로도 가치를 지닌다. 가령 실시간으로 그려지는 뇌파 그래프에 따라 이미지도 실시간으로 그리도록 구현한다면 시각적으로 주목할 만한 매체가 될 것이다.

# 과제 수행 방법

## 개발 환경

* + - * 예시) 활용 장비: GPU 서버 (대여)
      * 예시) 웹 브라우저 (예: Chrome, Firefox) - SQL 공격 시뮬레이션 웹 사이트
      * 예시) SQL Injection 공격 도구: Python

## 조사 방법론

* + - * 예시) 문헌 조사: 학술 논문, 기술 보고서, 관련 기관 자료 등을 활용하여 SQL Injection 공격 기법, 방어 전략, 기존 연구 등에 대한 정보를 수집한다.
      * 예시) 전문가 인터뷰: 보안 전문가, 웹 개발자 등을 인터뷰하여 실제 공격 사례, 현장에서의 문제점, 해결 방안 등에 대한 정보를 얻는다.
      * 실험 및 분석: 실제 모의 공격용 웹 애플리케이션을 대상으로 SQL Injection 공격 시뮬레이션을 수행하고 취약점을 분석한다.
      * 예시) bound box regression, image classification, object detection, semi-supervised learning
      * 예시) 활용 모델: YOLO

## 팀원 구성 및 역할

|  |  |  |  |
| --- | --- | --- | --- |
|  | 이름 | 역할 | 세부 역할 ( 프로젝트 개발 업무) |
| 1 | 000 | 팀장 | 예시) 프로젝트 조정 관리, 백엔드 개발 |
| 2 | 000 | 팀원 | 예시) 프론트엔드 개발 |
| 3 | 000 | 팀원 | 예시) 시스템 설계 |
| 4 | 000 | 팀원 | 예시) 네트워킹 기능 구현 |

# 과제 수행 일정

## 수행 계획

### 예시) 기술 현황 조사

* + - * SQL Injection 공격 기법 (Union-based, Boolean-based, Error-based 등)에 대한 심층 분석을 진행한다. 이때 기존 SQL Injection 방어 전략 및 알고리즘 (예: 입력 유효성 검사, 파라미터화 된 쿼리)의 장단점 역시 비교 분석한다.

### 시나리오 설계

* + - * 다양한 웹 애플리케이션 유형 (예: 전자 상거래, 소셜 네트워킹)을 고려한 공격 시나리오를 설계한다. 공격자의 목표 (데이터 탈취, 시스템 손상 등)에 따른 공격 시나리오를 구분하며, 공격 성공 시 발생 가능한 피해 규모 및 심각성을 분석해 기술한다

### 알고리즘 설계

* + - * SQL Injection 취약점 탐지 및 방어를 위한 새로운 알고리즘을 개발한다. 이때 알고리즘의 작동 원리, 장점 및 기존 방법 과의 차별성을 명확하게 제시하며, 웹 애플리케이션에 효과적으로 적용할 수 있도록 설계한다.

### 시스템 구현

* + - * 개발된 알고리즘을 실제 웹 애플리케이션에 적용하여 구현해본다. 다양한 공격 시나리오에 대한 실험을 통해 성능, 정확도 및 효과를 평가하며, 실험 결과를 표, 그래프 등을 활용하여 시각적으로 제시한다.

### 기능 검증 및 개선

* + - * 알고리즘의 기능 및 성능을 검증하고 문제점을 분석한다. 해당 알고리즘을 이용하며 2.2에서 설계한 공격 시나리오로 모의 공격을 진행하고, 발견된 문제점을 개선한다.

### 최종 성능 검증

* + - * 2.2에서 설계한 다양한 공격 시나리오에 대한 알고리즘의 성능을 최종적으로 검증해 목표치와 비교한다.
      * 공격 탐지 율, 정확도, 시스템 오버헤드 등을 측정하며, 검증 결과를 보고서에 표, 그래프 등을 활용해 시각적으로 상세히 제시한다.

### 보고서 작성 및 발표

## 세부 일정

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 추진내용 | Month1 | Month2 | Month3 | Month3 |
| 1 | 예시) 기술 현황 조사 |  |  |  |  |
| 2 | 공격 시나리오 설계 |  |  |  |  |
| 3 | 알고리즘 개발 및 설계 |  |  |  |  |
| 4 | 시스템 구현 및 실험 |  |  |  |  |
| 5 | 중간 검증 및 기능 개선 |  |  |  |  |
| 6 | 최종 성능 검증 |  |  |  |  |
| 7 | 보고서 작성 및 발표 |  |  |  |  |