

성균관대학교 – 소이넷

- 딥러닝 모델의 SoyNet 포팅 -

2023. 06. 13

성균관대학교 소프트웨어융합대학

발표자: 백승렬 (소프트웨어학과, 2학년)

2023 산학협력 과제 요약

과제명	딥러닝 모델의 SoyNet 포팅	기업명/ 담당자/ 멘토	소이넷 / 유경석 팀장님 /
과제개요	<ul style="list-style-type: none"> ● 딥러닝 모델들 분석 및 Config 파일 작성 ● 딥러닝 모델의 weight 초기값 변경 ● C++을 이용해 딥러닝 모델 최적화 ● 다수의 딥러닝 모델들 SoyNet에 포팅 	참여 학생	김호재 (소프트웨어학과, 2학년)
			백승렬 (소프트웨어학과, 2학년)
			박제현 (소프트웨어학과, 2학년)
			박세훈 (소프트웨어학과, 2학년)
			이해성 (소프트웨어학과, 2학년)
과제 기간	2023년 4월 ~ 2023년 12월 (9개월)	지도교수	이미향 교수
과제 최종 결과물	<ul style="list-style-type: none"> • SoyNet에 포팅된 최적화된 딥러닝 모델들 		

목차

- 팀 소개
- 기업 소개
- 기업의 Needs 분석
- 과제 내용
- 과제 Roadmap

팀 소개



김호재



백승렬



이미향 교수님



박세훈



박제현



이해성

기업 소개

소이넷은 C/C++ 기반의 **추론 가속화** 소프트웨어를 제공합니다.

- **추론 가속화 모델 종류**

Yolov5 s, IDN, Yolor 등의 모델들을 제공하고 있습니다. (참고: <https://soynet.io/market/>)

- **추론 비용 절감**

추론 과정이 3배 더 빨라지고, 사용되는 메모리를 9배 이상 절약합니다. 자원을 아껴 서버 비용을 절감합니다.

- **호환성**




TensorFlow / PyTorch / Caffe 등의 프레임 워크로 작성된 AI 모델도 정상적으로 지원합니다.

- **상품**

Free Subscription, Standard Support, Enterprise Support 상품을 제공합니다. (참고: <https://soynet.io/pricing/>)

기업 소개

소이넷은 여러 기업들과 함께 하고 있습니다.

		
POSCO <small>customer</small>	Hyundai Steel <small>customer</small>	Neuromeka <small>customer</small>
Surface Detection	Robotic Arm Setup	Neuromeka

기업의 Needs 분석

□ 주제 선정 동기

소이넷에서 만든 딥러닝 추론 솔루션 **SoyNet**은 딥러닝모델의 **추론과정에서 속도를 향상하고 경량화를 시켜줄 수 있는** 솔루션이다. 다만, 각각의 **딥러닝 모델은 구현된 방식이 서로 다르기 때문에** 소이넷의 추론 솔루션을 적용하기 위해서는 각각의 딥러닝 모델에 맞추어 **솔루션을 직접 포팅하여** 솔루션을 적용하는 **최적화 작업**을 해야한다.

□ 결과물에 대한 활용 계획

SoyNet의 추론 솔루션에 포팅한 각각의 딥러닝 모델들을 통해 기존에 비해 학습 생성된 모델들의 실제로 사용이 되는 추론과정에서의 메모리 효율성과 속도를 높여 딥러닝 모델들의 **편의성과 경제성을 향상할** 예정이다.

과제 내용

□ 과제 목표

SoyNet 추론 솔루션을 딥러닝 모델에 적용하는 포팅작업을 통해 **딥러닝 모델의 속도 향상과 메모리 효율성 증대.**

□ 선행기술

- C / C++
- 머신러닝과 딥러닝의 구조와 원리 이해
- 멀티쓰레드와 멀티프로세스에 관한 이해.
- Cuda / TensorRT programming
- PyTorch / Tensorflow

과제 내용

□ 개발 내용

- 각각의 딥러닝 모델들을 분석
- SoyNet의 추론 솔루션 기준에 맞추어 각 딥러닝 모델들에 적용되는 **Weight의 초기값 변경 및 config 값 변경**
- 파이썬으로 작성된 코드를 최적화 하기위해 그에 적합한 **C/C++ 코드로 재작성하여 변경**

□ 예상 결과물

SoyNet에 포팅된 최적화된 딥러닝 모델들

산학 협력 프로젝트 로드맵

팀	1학기	하계근무 집중교육기간 (6/12 ~ 6/25)	하계 근무 기간 (6/26 ~ 8/4)	2학기
선행기술 학습	ML/DL 학습 프로그래밍 기술 실습			
개발 (소이넷 포팅작업)			딥러닝 모델들의 소이넷 포팅 작업	
결과 정리				결과 정리 보고서 작성