

Team ALFA

2024 전기 졸업과제 착수보고서

강화학습기반 교차로 동적 신호제어시스템

권오성 이준표 정하림

목차

배경	2
목표	3
아키텍처 설계	4
사용기술 소개	5
관련 연구의 강화학습 모델	7
진행계획	8
제약사항	10
개발일정 및 역할분담	11
참고자료	12

배경

1. 횡단보도에서의 불필요한 대기 시간 문제

교차로에서 보행자 신호의 대기 시간이 필요 이상으로 길거나, 통행하는 차량 수가 적은 경우에도 횡단 신호등이 적색인 경우가 종종 발생합니다. 이로 인해 보행자는 지나치게 긴 대기 시간으로 인한 피로감을 느끼게 되며, 무단 횡단을 유발하는 원인이 됩니다. 반대로 차량 운전자의 입장에서 보면, 야간이나 새벽에 보행자가 거의 없음에도 불구하고 횡단보도 앞에서 오래 기다려야 하는 경우가 있습니다.

2. 교통 약자를 위한 적절한 횡단보도 신호 시간 필요성

보행자가 횡단 신호가 녹색으로 바뀐 직후에 횡단보도에 도착하거나, 보행 속도가 느린 어린이나 노약자 등의 경우, 제 시간에 횡단보도를 건너기 어려운 경우가 많습니다. 이들이 신호가 바뀐 후에도 위험하게 도로를 벗어나는 경우를 종종 볼 수 있습니다. 횡단보도의 길이에 비해 짧은 녹색 신호 시간을 주는 신호등도 문제입니다.

3. 현재 신호 체계의 효율성과 유연성 부족

현재 국내에서 사용되는 신호 체계는 특정 상황별로 지정된 주기를 준비해 두고 그 상황이 발생하면 적용하는 시간 제어식 신호를 주로 사용하고 있습니다. 하지만 이 방식은 교통량의 변동에 따라 해당 구간의 신호에 대한 인위적인 조작이 필요하므로, 효율성이 떨어지며 빠른 대응이 어렵고 여러 구간에 걸친 미세 조정이 힘듭니다. 또한 횡단 보행자 요소를 고려하지 않는 것도 문제입니다. 따라서, 인구가 몰리고 차량 대수가 늘어나는 현대 도시에서는 불확실성이 증가하므로, 보다 동적이고 유연한 신호 체계를 위해 강화학습을 활용할 필요가 있습니다.

목표

1. 횡단보도 앞 불필요한 대기시간을 줄이기 위한 동적 신호체계

교차로 도로변에 카메라를 설치하여 주변 도로와 보행자를 실시간으로 인식하고, 횡단보도에 사람이 (다수 혹은 오래) 기다리고 있으나 도로에 지나다니는 차량이 적을 경우, 횡단 신호를 변경하거나 신호 길이/주기를 조절하는 등의 동적인 신호 체계를 설계합니다. 동시에 차량 운전자의 입장에서도 신호를 최적화하도록 합니다.

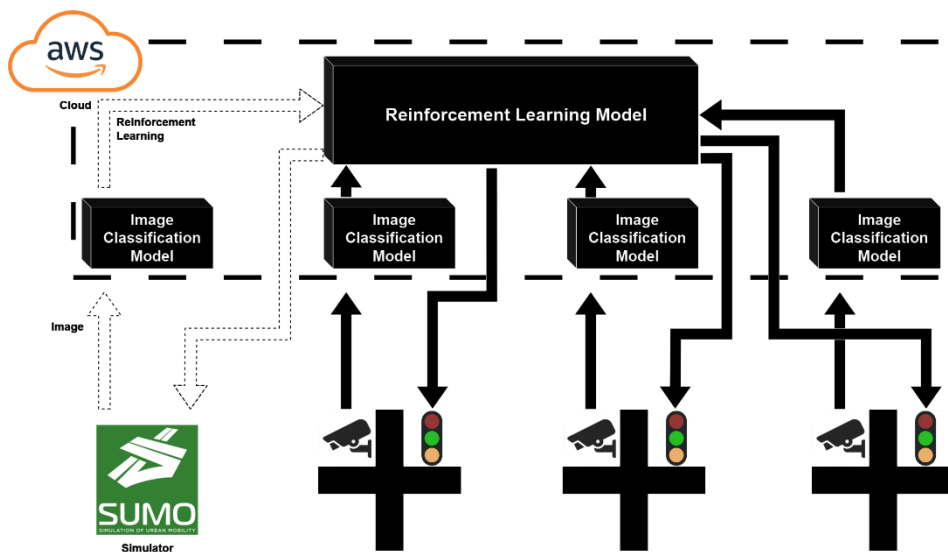
2. 보행자가 안전하게 횡단보도를 건너도록 신호길이를 조절하는 동적 신호체계

보행 속도가 느린 어린이나 노약자 등의 보행자들을 위해, 이들이 안전하게 건널 수 있도록 녹색 신호 시간을 상황에 맞게 동적으로 늘려주는 시스템을 설계합니다. 교차로의 카메라로 횡단보도 주변에서 이동 중인 보행자들의 속도와 행태 등을 실시간으로 분석하고, 필요시 연계된 신호등의 신호 시간을 적정량 증가시킵니다.

3. 교통량과 보행인파 모두를 고려한 실시간 AI 최적화 도시교통신호체계

보다 효율적이고 유연한 교통 신호 체계를 구축하려면, 도로 구간의 자동차 통행은 물론, 각 교차로 지점의 횡단 보행자 요소도 함께 고려해야 합니다. 이를 위해 각 교차로에 설치된 카메라를 활용하여 실시간으로 위치 별 차량과 보행자를 인식하여 데이터를 수집합니다. 이렇게 수집한 데이터를 클라우드 환경에서 실시간으로 강화학습 모델로 추론한 후, 각 구간별 최적의 신호 상태를 도출하여 도시 전체의 신호 시스템에 적용하면, **거시적인 교통 로드 밸런싱**이 가능해질 것입니다.

아키텍처 설계



클라우드 환경(AWS or Azure)에서 교차로 상황(객체) 인식 모델과 강화학습 모델을 컨테이너로 구성하고 쿠버네티스를 이용하여 배포합니다. 실제 적용은 여러 교차로에서 촬영한 영상을 클라우드로 전송하고 상황(객체) 인식 모델이 상황을 분석하고 강화학습 모델이 결정을 내려 다시 교차로 신호시스템을 조작하는 구조입니다. 하지만 실제 교차로에 적용을 할 수 없기 때문에 시뮬레이터와 상호 작용하는 아키텍처를 택하여 구현할 예정입니다.

사용기술 소개

1. 상황(객체) 인식 모델

YOLO (You Only Look Once)는 객체 인식을 위한 실시간 방법론입니다. 이 방법론은 객체 인식을 공간적으로 분리된 경계 상자와 관련 클래스 확률에 대한 회귀 문제로 구성합니다. 전체 이미지에서 직접 경계 상자와 클래스 확률을 예측하는 단일 신경망을 사용합니다. 이로 인해 전체 탐지 파이프라인이 단일 네트워크가 되므로, 탐지 성능에 직접적으로 최적화할 수 있습니다. YOLO의 통합 아키텍처는 매우 빠르며, 기본 YOLO 모델은 초당 45 프레임으로 실시간으로 이미지를 처리합니다. 더 작은 버전의 네트워크인 Fast YOLO는 초당 놀라운 155 프레임을 처리하면서도 다른 실시간 탐지기의 mAP를 두 배로 늘립니다.

SSD (Single Shot MultiBox Detector)

Faster R-CNN

2. 강화학습 알고리즘

Q-Learning은 모델 없이 학습하는 강화학습 기법 중 하나입니다. 강화학습에서는 MDP(markov decision process)를 통해 학습을 모델링 합니다. 주어진 환경에서 상태를 인식하고 학습을 통해 최적의 행동을 선택합니다. 현재 상태에서 어떤 행동을 수행할지 선택하는 기준은 정책을 통해 결정됩니다. 선택된 행동으로 얻는 보상이 생기고, 보상을 바탕으로 경험을 쌓으며 다음 상태로 넘어갑니다.

DQN(Deep Q-Network)는 Q-Learning에 심층 신경망을 결합한 강화학습으로, Q값을 테이블이 아닌 심층 신경망을 이용해서 학습하는 기법입니다. Experience Replay를 통해 에이전트가 이전 경험을 재사용하여 학습할 수 있게 하여 학습 안정성을 향상시키고, Target network를 사용하여 학습과정 중 사용되는 Q값을 안정화하여 수렴을 더 빠르게 만듭니다. 더불어, 대규모의 상태 및 행동 공간에서도 효과적으로 학습할 수 있습니다.

3. 교통 시뮬레이터

SUMO는 도시 교통의 시뮬레이션을 위한 오픈 소스 소프트웨어입니다. 대규모 네트워크를 처리하기 위해 설계된 미시적이고 연속적인 다중 모드 교통 시뮬레이션 패키지로, 도로 차량, 대중교통, 보행자 등의 교통 시스템을 모델링할 수 있습니다. 또한 경로 찾기, 시각화, 네트워크 가져오기, 배출량 계산 등의 작업을 처리하는 다양한 지원 도구를 포함하고 있습니다. 사용자 정의 모델로 SUMO를 확장할 수 있으며, 시뮬레이션을 원격으로 제어하는 데 사용할 수 있는 다양한 API를 제공합니다.

VISSIM은 PTV 그룹에서 개발한 교통 시뮬레이션 소프트웨어입니다. 모든 도로 사용자의 교통 패턴을 미시적인 규모로 디지털로 재현합니다. 과학적으로 기반한 시뮬레이션과 시나리오 관리를 통해 교통 인프라의 성능을 평가하고 최적화하고, 데이터 기반의 계획 결정을 하며, 혼잡, 배출, 그리고 다른 모드에 대한 도로 공간의 공정한 분배와 같은 도전 과제를 미리 대응할 수 있습니다. 도로, 자전거, 또는 대중교통 인프라(철도 포함)뿐만 아니라 보행자 도로를 3D로 현실적으로 표현할 수 있습니다. 주차장, 공항, 기차역, 큰 건물 또는 임시 조치도 시뮬레이션하고 평가할 수 있습니다.

관련 연구의 강화학습 모델

주현진 (2021) 연구에서는 단일 교차로 환경과 다중 교차로 환경에서 교통 신호를 제어하는 두 가지 모델을 설계하였습니다. 하나는 녹색 신호 순서 할당 모델로, 이는 이동 신호를 결정하여 차량의 흐름을 제어합니다. 다른 하나는 녹색 신호 시간 할당 모델로, 이는 녹색 신호의 지속 시간을 결정합니다.

이 모델들은 강화학습을 사용하여 교차로의 상황을 인식하고, 다양한 매개변수들의 활용성과 유용성을 학습합니다. 이를 통해 차량의 흐름을 제어하고, 교차로에서의 처리량을 최대화하는 것을 목표로 합니다.

모델은 MDP(Markov decision process)를 사용하여 상태, 행동, 그리고 보상을 정의합니다. 상태는 교차로의 방향의 개수로 정의되며, 행동은 동시에 녹색 신호를 할당할 수 있는 방향의 집합으로 정의됩니다. 보상은 큐 길이의 표준편차와 처리량으로 구성됩니다.

큐 길이의 표준편차는 교차로에 있는 차선의 길이가 얼마나 흩어져 있는지를 나타내며, 처리량은 교차로를 지나간 차량의 개수를 나타냅니다. 이 두 가지 요소를 최적화함으로써, 교차로에서의 차량 흐름을 효율적으로 제어하고, 처리량을 최대화하는 것이 가능해집니다.

구자운 (2022) 연구에서는 상태, 행동, 보상을 다음과 같이 정의하였다.

상태는 각 차선에 정지한 차량의 수, 현재 신호의 등화 여부, 그리고 현재 등화된 신호의 경과 시간으로 정의됩니다. 이러한 정보는 에이전트가 어떤 행동을 취할지 결정하는 데 사용됩니다.

행동은 현재 신호를 유지할지, 아니면 다음 신호로 변경할지를 결정하는 것입니다. 이 모델은 신호의 순서를 유지하고, 각 신호는 한 번의 신호 주기 동안 최소 한 번은 등화되어야 한다는 제약 조건을 가지고 있습니다.

보상은 행동에 따라 변경된 교통 흐름을 모델에 제공합니다. 보상은 정지 차량 대비 통과 차량의 비율로 정의되며, 이를 통해 정지 차량이 줄고, 통과 차량이 많아질수록 에이전트가 받는 보상이 커지게 됩니다.

진행계획

1. 상황(객체) 인식 모델

VISSIM 시뮬레이터로 구현한 3D 교차로 영상에서 차량과 보행자 및 차선과 횡단보도 객체의 이미지를 수집하여 학습하고 검증합니다. 인식한 객체인 차량과 보행자의 속도, 대기시간, 종류 등 강화학습 모델에 필요한 파라미터들을 실시간으로 계산하여 출력할 수 있도록 합니다.

2. 강화학습 모델

SUMO 시뮬레이터에서 제공하는 TraCI 인터페이스를 사용하여 시뮬레이터-모델 간 상호작용하며 학습합니다. 기존 연구에서는 차량 $상대속도$ 만을 고려하여 최적화하는 방식을 택하고 있는데 이번 프로젝트에서는 보행자의 $상대속도$ 도 함께 고려하여 모델을 설계합니다. 즉 교차로 범위내에 대기하고 있는 차량과 보행자의 수로 정의합니다.

$행동$ 은 기존 연구에서는 신호 순서와 신호 시간을 분리하여 모델을 만들고 비교하였는데 이번 프로젝트에서는 두가지 요소 모두를 고려하여 설계합니다. 교차로에는 가능한 신호 패턴이 정해져 있고 에이전트는 현재 녹색 신호 시간과 다음 신호 패턴을 결정합니다.

$보상$ 은 불쾌지수를 도입하여 교차로에 대기 중인 모든 객체(차량, 보행자)의 불쾌지수가 최소화될 수 있도록 합니다. 불쾌지수는 객체의 대기 시간에 비례하여 지수적으로 증가하도록 설계합니다. 기존 연구에서는 큐의 길이와 대기 시간을 최소화하는 방법을 택하였는데 불쾌지수를 도입하게 되면 큐의 길이와 대기 시간을 종합적으로 고려하는 모델이 됩니다. 왜냐하면 큐의 길이가 길어지면 모든 객체의 불쾌지수를 합하기 때문에 불쾌지수의 총량이 증가하고 대기 시간이 길어지면 각 객체의 불쾌지수가 상승하기 때문에 불쾌지수의 총량이 증가하게 됩니다. 또한 객체별 가중치를 다르게 하여 버스와 같은 대중교통 차량이 다른 객체에 비해서 빠르게 통과할 수 있도록 합니다.

3. 상황(객체) 인식 모델과 강화학습 모델의 통합

강화학습 모델에서 필요한 파라미터들을 상황(객체) 인식 모델이 정확히 제공해 줄 수 있어야 합니다. 두 모델이 다른 시뮬레이터로 학습을 하였기 때문에 원활한 상호작용을 위해 from 상황(객체) 인식 모델 to 강화학습 모델 인터페이스와 from 강화학습 모델 to VISSIM 시뮬레이터 인터페이스를 개발합니다.

4. 최종 결과물

교통신호 관리자가 실시간으로 교차로의 상황(교차로 영상, 신호 상태 등)을 알 수 있도록 하고, 이상 징후를 판단할 수 있는 metric 을 제공합니다. 이상 징후 발생시 실시간으로 각 교차로의 신호를 제어할 수 있도록 웹 어플리케이션 혹은 프로그램을 구현합니다.

제약사항

1. 데이터 다양성의 부족

최초 객체 인식 학습 시에 사용할 온라인에서 구한 이미지 데이터들의 소스가 다양하지 않으면, 여러 각도와 다양한 날씨 및 명암 환경에서의 교차로 객체들을 인식하는 데 한계가 있을 수 있습니다. 이는 모델의 일반화 능력을 제한할 수 있습니다.

2. 시뮬레이션과 현실 간의 괴리

모델 검증 시 실제 교차로에서 차량과 보행자들에 대해 테스트하는 것이 아니라 시뮬레이션을 통해 검증하므로, 현실의 결과와 괴리가 발생할 수 있습니다. 이는 모델의 실용성을 제한할 수 있습니다.

3. 모델의 규모와 리소스의 한계

도시 규모의 교통 신호 체계를 상정했지만, 실제로는 모델의 사이즈나 기술, 리소스의 한계로 인해 비교적 적은 수의 교차로(예: 1~3 개)에 대해서만 연구를 진행할 수 있을 것으로 보입니다. 이 경우 실제 도시 규모인 수십~수백 개의 교차로 환경에서는 어떤 결과가 나올지 예측하기 어렵습니다.

4. 네트워크 환경의 제약

각 교차로의 엣지에서 수집한 방대한 영상 데이터를 실시간으로 중앙 서버에 전송하기 위한 네트워크 환경 구현에 어려움이 있을 수 있습니다. 이는 데이터 처리와 모델 학습의 효율성을 제한할 수 있습니다.

개발일정 및 역할분담

일정	7 월				8 월				9 월				10 월			
	1 주	2 주	3 주	4 주	1 주	2 주	3 주	4 주	1 주	2 주	3 주	4 주	1 주	2 주	3 주	4 주
평가지표 설계																
시뮬레이터 테스트																
강화학습 모델 설계																
강화학습 모델 학습																
강화학습 모델 검증																
객체 인식 모델 설계																
객체 인식 모델 학습																
객체 인식 모델 검증																
모델 통합 테스트																
클라우드 환경 구축																
통합 테스트 환경 구축																
통합 테스트																

권오성: 평가지표 설계, 강화학습 모델, 클라우드 환경 구축, 통합 테스트 환경 구축

이준표: 시뮬레이터(VISSIM) 테스트, 객체 인식 모델, 통합 테스트

정하림: 시뮬레이터(SUMO) 테스트, 객체 인식 모델, 모델 통합 테스트

참고자료

구자운 (2022) 강화학습을 이용한 다중 교차로에서의 신호 최적화

주현진 (2021) 스마트 도시에서 강화학습 기반 지능형 교통신호 시스템 최적화 연구

도철웅, 최기주, 오철 저 교통공학원론 제 3 개정판 (주)교문사