

# FCN-rLSTM 모델에 기반한 교통량 분석 시스템

권태윤<sup>01</sup> 윤찬웅<sup>2</sup> 손홍일<sup>3</sup> 김기태<sup>4</sup> Scott Uk-Jin Lee<sup>5</sup>

<sup>1,2,3,4,5</sup> 한양대학교 ERICA 소프트웨어융합대학 소프트웨어학부

<sup>1,2,3,4,5</sup>{kwonconnor101, qicqockqvvgiq, hongil7626, kimkt0155}@gmail.com,

<sup>5</sup>scottlee@hanyang.ac.kr

## Traffic Analysis System based on FCN-rLSTM model

Taeyoon Kwon<sup>01</sup>, Chanwoong Yoon<sup>2</sup>, Hongil Son<sup>3</sup>, Gitae Kim<sup>4</sup>, Scott Uk-Jin Lee<sup>5</sup>

<sup>1,2,3,4,5</sup> Division of Computer Science, College of Computing, Hanyang University ERICA

### 요 약

지능형교통체계(ITS)를 구축하기 위해서 교통 정보의 실시간 수집 및 분석은 필수적으로 요구된다. 본 논문에서는 이러한 교통 정보의 실시간 수집 및 분석을 효율적이고 효과적으로 수행하기 위해 차량 계수 예측 모델 중 state-of-the-art 모델인 FCN-rLSTM 모델을 활용하여 다수의 실시간 도로 CCTV 영상을 병렬적으로 분석하고 또한, 사용자의 요청에 따라 분석 결과를 제공해주는 통합 교통량 분석 시스템을 제안한다.

### 1. 서 론

지능형교통체계(ITS)는 교통수단의 실시간 관리, 제어와 교통 정보의 실시간 수집 및 활용을 통해 시민들의 교통편의 증진과 교통안전을 도모할 수 있도록 교통체계의 운영, 관리를 자동화 및 과학화하는 환경 친화적 미래형 교통체계이다[1]. 교통 정보의 실시간 수집 및 분석은 지능형교통체계를 구축하기 위해 요구되는 기본적인 구성요소이기에 이를 효율적이고 효과적으로 처리하기 위한 다양한 연구들이 수행되고 있다.

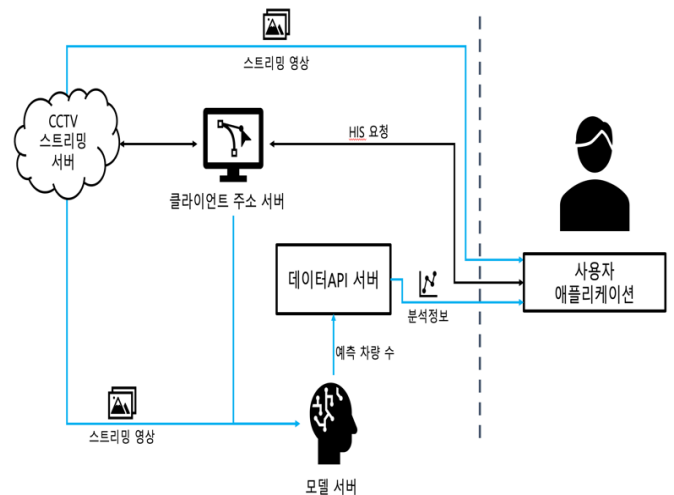
본 논문에서는 컴퓨터 비전 분야의 차량 계수 예측 모델 중 state-of-the-art 모델인 FCN-rLSTM[2] 모델을 활용하여 다수의 실시간 CCTV 영상을 병렬적으로 처리함을 통해 효율적으로 교통정보를 수집하고, 수집한 정보를 실시간으로 제공하거나 이를 기간별로 분석한 결과 및 통합 데이터를 제공하여 분석된 정보를 효과적으로 사용할 수 있도록 하는 시스템에 대해 제안한다.

시스템의 교통량 분석에 사용되는 FCN-rLSTM 모델은 fully convolutional neural networks(FCN)와 long short term memory networks(LSTM)를 잔차를 학습하는 방식을 통해 결합하여 차량의 개수를 예측하는 모델이며, FCN 계층에 의한 픽셀 수준 예측의 장점과 LSTM 계층에 의한 시간 역학적 정보를 학습할 수 있는 장점을 결합한 차량 수 예측 모델이다[2].

### 2. 교통량 분석 시스템<sup>1</sup>

#### 2.1 시스템 개요

본 시스템은 [그림 1]과 같은 개요를 가진다. CCTV 스트리밍 서버는 국가교통정보에서 제공하는 대한민국



[그림 1] 시스템 전체 개요도

실시간 교통 CCTV 영상[3]을 사용했으며, 이는 HTTP 라이브 스트리밍(hls)을 통해 실시간 영상을 제공한다. 이러한 영상은 사용자의 실시간 분석 작업과 모델 서버에서의 실시간 데이터 수집 작업을 진행하는 경우에 사용이 된다. 모델 서버에서는 클라이언트 주소 서버에서 제공받은 CCTV 스트리밍 서버의 hls 주소 정보를 통해 실시간 영상을 받아오게 된다. 이를 1 fps로 프레임을 추출한 후, FCN-rLSTM[2] 모델을 사용하여 영상 속 차량 수를 예측하며 모델 예측 데이터는 소켓 통신에 기반하여 데이터 API 서버로 1초 간격으로 전달한다. 데이터 API 서버에서는 전달받은 데이터를 CCTV 마다 배열에 저장하며, 1시간 단위로 파일 시스템에 CSV 파일 형태로 저장시키도록 설계를 하였으며 또한, 사용자의 요청에 따라 특정 CCTV의 실시간 분석 결과 혹은, 기간별 분석 데이터를 사용자에게 제공해주는 기능을 가

<sup>1</sup>본 연구는 2022년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학지원사업의 연구결과로 수행되었음(2018-0-00192)

[표 1] 학습 모델과 FCN-rLSTM[2]모델의 MAE 비교

	WebCamT[4]	Trancos[5]
FCN-rLSTM[2]	1.53	4.21
학습 모델	1.61	4.28

진다. 사용자 애플리케이션에서는 실시간 분석 영상을 보여주는 기능과 기간별 분석 데이터를 보여주는 기능을 제공하도록 설계를 하였다.

## 2.2 클라이언트 주소 서버

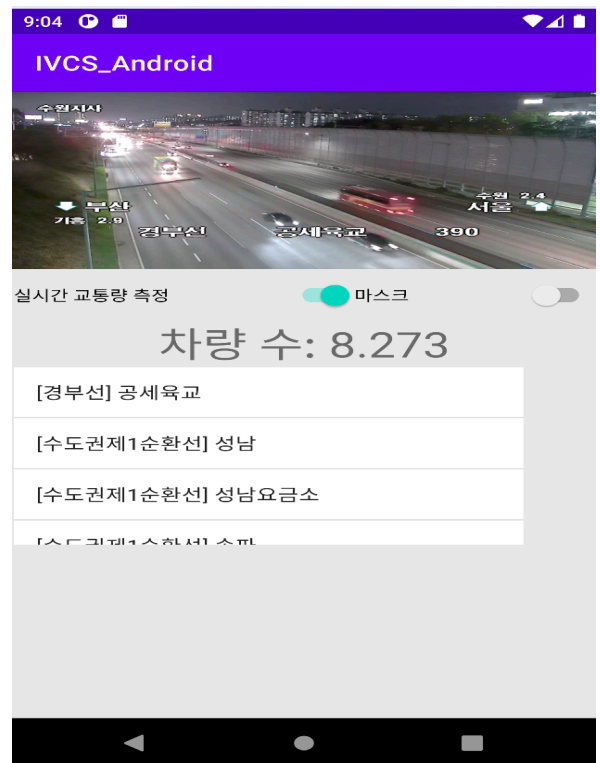
클라이언트 주소 서버에서의 클라이언트는 시스템을 사용하는 사용자와 모델 서버에 해당되며 스트리밍을 위한 hls 주소를 보유하고 배포하는 역할을 한다. 해당 서버는 처음 구동할 때와 hls 주소가 만료 되었을 경우 CCTV 스트리밍 서버에 hls 주소를 요청함을 통해 주소 정보를 지속적으로 업데이트한다. 그리하여, 클라이언트의 요청에 따라 보유하고 있는 hls 주소를 HTTP GET 메소드를 통해 JSON 형식으로 전송하게 된다.

## 2.3 FCN-rLSTM 모델 및 모델 서버

시스템의 모델 서버에서 사용되는 FCN-rLSTM[2] 모델은 FCN 계층에서는 픽셀 수준의 차량 밀집도를 예측하게 되며 차량 밀집도에 있는 모든 픽셀 값을 합하면 차량 수를 얻을 수 있다. 이렇게 FCN 계층에서 얻은 여러 개의 차량 밀집도를 LSTM 계층의 입력으로 사용하게 되며, LSTM 계층에서는 이전 이미지의 차량 밀집도 정보를 기억하여 현재 이미지의 차량 밀집도 정보를 이용한 차량 수 추론에 도움을 주게 된다. 이 모델은 대규모 교통 데이터 셋인 WebCamT[4] 에서 평균 절대 오차(MAE)가 1.53 인 성능을 보여주며, 또 다른 교통 데이터 셋인 TRANCOS[5]에서 MAE 가 4.21 인 성능을 보이게 된다[2].

제안한 시스템의 모델 서버에서 해당 모델을 사용하기 위해 FCN-rLSTM 논문[2]에 입각하여 학습을 진행하였으며, 학습을 위한 설정 또한 동일하게 진행하였다. 학습은 WebCamT[4] 데이터셋을 사용하였으며, 성능은 [표 1]과 같이 평균절대오차(MAE)가 WebCamT[4] 데이터셋에서 1.61, TRANCOS[5] 데이터셋에서 4.28 로 FCN-rLSTM[2] 모델과 유사한 성능을 가지는 것으로 판단하여 본 시스템에 적용하였다.

제안한 시스템의 모델 서버에서는 실시간 CCTV 영상을 받기 위해 클라이언트 서버에서 CCTV 스트리밍 서버의 hls 주소 정보를 받아오게 된다. 이를 통해 hls 프로토콜로 영상을 받아오게 되며 1 fps로 프레임을 추출하여 모델의 입력으로 사용할 이미지를 얻게 된다. 프레임을 추출할 때 시스템상의 지연이 발생하게 되면, 지연된 프레임을 건너뛰고 가장 최근의 프레임을 받아오게 함으로써 실시간 분석에 지장이 없게 설계가 되었다.



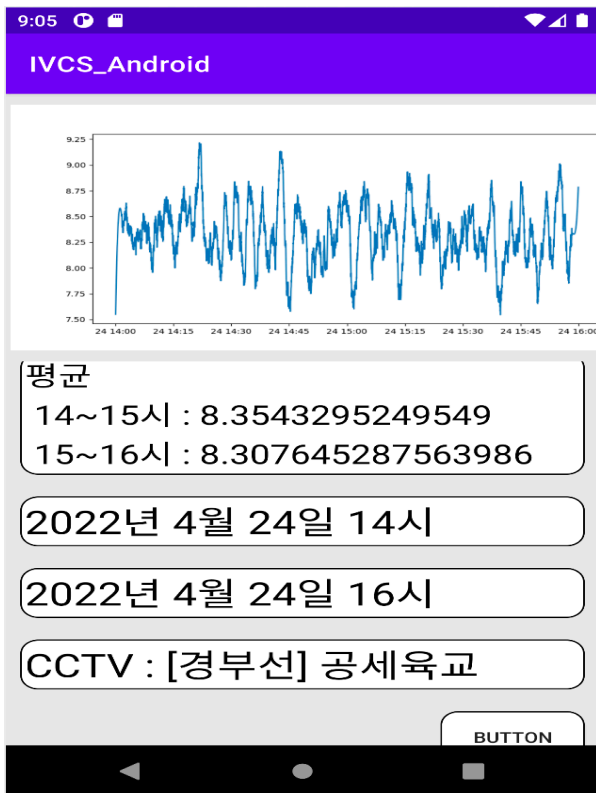
[그림 2] '[경부선] 공세육교' CCTV의 사용자 애플리케이션 실시간 분석 실행 화면

FCN-rLSTM[2] 모델은 LSTM 구조를 사용하기에 연속된 이미지 배열을 입력으로 요구한다. 모델은 배열의 크기가 5일때 가장 뛰어난 성능을 보이기에[2] CCTV 영상에서 추출한 실시간 프레임의 크기가 5인 큐 배열에 넣게 되며, 이렇게 얻은 큐 배열을 모델의 입력 요소로 사용하게 된다. 이 와 같은 과정을 여러 CCTV 영상에 동시 적용함으로써 크기가 5인 큐가 여러 개 겹친 텐서를 모델의 입력으로 사용하며 모델을 통해 CCTV 개수와 동일한 출력 결과의 개수를 얻게 된다. 이렇게 얻게 된 데이터는 소켓 통신 기반으로 데이터 API 서버로 전달하게 된다.

## 2.4 데이터 API 서버

데이터 API 서버에서는 두 가지 작업을 수행한다. 한 가지는 모델 서버에서 전달 받은 데이터를 파일 시스템을 이용하여 {CCTV 이름/ 연도/ 월/ 일} 순서대로 세부 폴더에 나누어 CSV 파일로 저장하는 작업을 진행하며, 다른 한 가지 작업은 사용자가 특정 CCTV의 실시간 분석 결과 혹은 기간에 대한 분석 데이터를 요청 시 제공해주는 작업을 진행한다.

모델의 예측 데이터를 실시간으로 저장하기 위해서는 먼저 각 CCTV 마다 교통량 데이터를 저장하는 배열을 필요로 한다. 각 배열에는 (시간, 차량 수 예측 결과) 데이터를 저장하며 1시간 마다 CSV 파일로 저장한다. 본 서버는 모델 서버에서 예측 결과를 소켓 통신 기반



[그림 3] ‘[경부선] 공세육교’ CCTV 의 2022.04.25 14~16시 데이터 분석 실행 화면

으로 1초 마다 전달 받기에 저장 과정을 별도의 스텝을 이용해 비동기적으로 처리함으로써 실시간 데이터 수집 작업에 문제가 발생하지 않게 한다.

사용자가 특정 CCTV에 대한 실시간 분석 결과를 요청하는 경우에는 모델 서버에서 들어오는 예측 결과를 사용자에게 소켓 통신 기반으로 즉시 전달한다. 사용자가 기간별 분석 데이터를 요구하는 경우에는 분석 과정을 별도의 스텝을 통해 진행하여 시간-차량 수 그래프와 시간 당 평균 차량 수를 HTTP 프로토콜 기반으로 전달하게 된다.

## 2.5 사용자 애플리케이션

사용자 애플리케이션에서는 선택한 CCTV 에 대한 실시간 영상 및 예측 결과를 보여주는 기능과 기간별 분석 데이터를 보여주는 기능을 보여준다.

실시간 분석 영상은 원하는 CCTV를 선택하면 CCTV 스트리밍 서버에서 hls 프로토콜을 통해 받아온 실시간 영상과 데이터 API 서버에서 받아온 예측 결과를 [그림 2]와 같이 나타내게 된다. 이 기능을 통해 사용자는 실시간으로 교통량을 수치화 하여 확인할 수 있게 된다.

기간별 특성 데이터는 특정 CCTV 와 기간을 선택하게 된다면 데이터 API 서버에 HTTP 통신 기반으로 요청하게 되며 해당 기간의 시간-차량 수 그래프와 시간 별 평균 차량 수 데이터를 전달받게 된다. 이렇게 전달 받은 데이터는 [그림 3]과 같이 사용자 애플리케이션에

서 보여지게 된다. 이를 통해 특정 기간의 교통량을 파악 및 분석할 수 있으며 이를 활용하면 지능형 교통체계를 구축하는데 많은 도움을 줄 것으로 판단된다.

## 3. 결론 및 향후 계획

본 논문에서는 다수의 실시간 도로 CCTV 영상을 병렬적으로 분석을 진행하며, 사용자의 요청에 따라 실시간 분석 결과 혹은 기간별 분석 데이터를 제공하는 교통량 분석 시스템을 제안하였다. 제안한 시스템을 활용한 실시간 교통 정보의 수집 및 활용은 지능형교통체계를 구축하는 데에 있어 이점을 줄 수 있을 것으로 판단된다.

향후 계획으로는 본 교통량 분석 시스템을 통해 교통량 데이터가 충분히 쌓이게 된다면 보다 다각적인 분석을 통해 현재로서는 부족한 분석 결과를 추가로 제공할 예정이며, 이는 지능형교통체계를 구축하는 데에 필수적 교통 정보의 분석에 많은 도움을 줄 것이라 판단된다.

## 4. 참고 자료 및 문헌

- [1] 국토교통부. [https://www.molit.go.kr/USR/policyData/m\\_34681/dtl?id=406](https://www.molit.go.kr/USR/policyData/m_34681/dtl?id=406)
- [2] Shanghang Zhang, Guanhang Wu, Joao P Costeira, and Jose MF Moura. Fcn-rlstm: Deep spatio-temporal neural networks for vehicle counting in city cameras. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3667–3676, 2017.
- [3] 국가교통정보. <https://www.its.go.kr/map/cctv>
- [4] Shanghang Zhang, Guanhang Wu, Joao P Costeira, and Jose MF Moura. Understanding traffic density from large-scale web camera data. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 5898–5907. 2017.
- [5] Daniel Onoro-Rubio, Roberto J. Lopez-Sastre. Towards perspective-free object counting with deeplearning. In European Conference on Computer Vision, pages 615–629. Springer, 2016