# A Bigdata Acquisition Framework of Deep-Learning-based CCTV-Video Contextualization Machines


Eun-Bee Cho*
eunbee0508@kgu.ac.kr
Data and Process Engineering Research Lab.
Department of Public Safety Bigdata, Graudate School of KYONGGI UNIVERSITY
Suwon-si, Gyeonggi-do, South Korea

Kyung-Hee Sun†
sunkyunghee@kgu.ac.kr
Contents Convergence Software Research Institute
KYONGGI UNIVERSITY
Suwon-si, Gyeonggi-do, South Korea

Dinh-Lam Pham‡
phamdinhlam@kgu.ac.kr
Contents Convergence Software Research Institute
KYONGGI UNIVERSITY
Suwon-si, Gyeonggi-do, South Korea

Kyoung-Sook Kim§
khmjmc@kgu.ac.kr
Contents Convergence Software Research Institute
KYONGGI UNIVERSITY
Suwon-si, Gyeonggi-do, South Korea

Jeong-Hyun Chang¶
crime_tiger564@kgu.ac.kr
Contents Convergence Software Research Institute
KYONGGI UNIVERSITY
Suwon-si, Gyeonggi-do, South Korea

Kwanghoon Pio Kim‖
kwang@kgu.ac.kr
Data and Process Engineering Research Lab.
Division of AI Computer Science and Engineering, KYONGGI UNIVERSITY
Suwon-si, Gyeonggi-do, South Korea


## 1 INTRODUCTION

In this paper, we initiate an innovative Bigdata acquisition approach called as *deep-learning-based CCTV-video contextualization machines*, which can scientifically and systematically obtain the innovative surveillance data-stores of active CCTV-contextual clues—objects, motions, and environs—extracted from streaming the surveillance CCTV-video devices by using the neural network architectures and their deep learning frameworks and models. The proposed framework is for fulfilling the full range of deep learning driven CCTV-contextual clue acquisition functionality that spans from the very early stage of building an architectural framework to the very late stage of applying the implemented framework to the so-called intelligent-led crime prediction and prevention as application services. As a tangible proof of the conceptual framework, we successfully implemented a CCTV-video object contextualization system that is supported by the YOLO (You Only Look Once) video-object detection neural network architecture and its variants [1][2][3].

## 2 FRAMEWORK, SYSTEM, AND VALIDATION

The essential component of the framework is the CCTV-video contextualization machine supported by the novel and innovative concept of the CCTV-video contextual clue detection approaches based upon the deep neural network models. The functional components of the machine are as follows: first, the CCTV-video manager is identifying video-frames from the input video clippings and files, second, the YOLO video-object detector is detecting YOLO-objects on each of the CCTV-video-frames, and third, the transformer is contextualizing the detected YOLO-objects and their properties into the context-objects represented in a textual formation of XML-schema and JSON formats. Finally, all the contextualized CCTV-video's YOLO-objects are transferred and stored on a cloud-based archive under the name of the active video-contexts Bigdata.

The eventual output of the CCTV-video-object contextualization machine is a contextual clue dataset of a corresponding CCTV-video clip, each of which is coded as COME-Code[1] in an XML schema structure. The COME-Code is for formatting the detected video-objects into the corresponding contextualized video-objects on all the video-frames of the CCTV-video clippings. Based upon the concrete framework, we implemented the CCTV-video contextualization system basically supporting the CCTV-video contextual clue (YOLO-object) detection functionality and the transformation functionality, as well. We tried to verify the functional correctness of the system through applying to a sample file of CCTV-video clippings captured from a real CCTV device installed at a street.

## 3 CONCLUSION

This paper proposed a novel concept of the CCTV-video contextualization machines and its Bigdata acquisition framework and verified the functional correctness of the machines via an experimental verification. The proposed Bigdata acquisition framework and its implemented system are tangible and applicable as a meaningful tool for the CCTV-video surveillance platforms by successfully implementing the CCTV-video object contextualization machines with the cutting-edge deep learning approach, YOLO. Finally, we have confidence on this system's expansibility and applicability in many video-related Bigdata engineering platforms and services.

## REFERENCES


[1] Tanvir Ahmad, Yinglong Ma, Muhammad Yahya, Belal Ahmad, Shah Nazir, and Amin ul Haq. 2020. Object Detection through Modified YOLO Neural Network. *Scientific Programming* 2020 (2020), 8403262.

[2] Tausif Diwan, G. Anirudh, and Jitendra V. Tembhurne. 2023. Object detection using YOLO: challenges, architectural successors, datasets and applications. *Multimedia Tools and Applications* 82 (March 2023), 9243–9275.

[3] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You Only Look Once: Unified, Real-Time Object Detection. *arXiv: 1506.02640v5 [cs.CV]* 5 (May 2016), 1–10.


[1] COME-Code stands for Contextualized Object, Motion, and Environ Code.