Face Recognition Using Faster R-CNN with Inception-V2 Architecture for CCTV Camera

Lavin J. Halawa
Department of Computer Science
Faculty of Mathematics and Science
Universitas Diponegoro
Semarang, Indonesia

Adi Wibowo
Department of Computer Science
Faculty of Mathematics and Science
Universitas Diponegoro
Semarang, Indonesia
bowo.adi@live.undip.ac.id

Ferda Ernawan
Faculty of Computing
College of Computing and
Applied Sciences
University Malaysia Pahang

Abstract— **Detection and prevention of criminal incidents using CCTV are currently increasing trend, for example, car and motorcycle parking lot. However, not continuous people monitoring and careless of events produce useless CCTV function for the prevention of criminal incidents. In this paper, face recognition is used for the recognition of vehicle owners in parking lots that are CCTV installed. The Faster-RCNN method is used for face detection and also for face recognition. Inception V2 architecture is utilized due to has a high accuracy among Convolutional Neural Network architecture. The best learning rate and epoch parameters for the Faster R-CNN model are optimized to improve face recognition on CCTV. In this research, the dataset consists of 6 people images with 50 faces images for each people, which used as training data, testing data, and validation data.**

Keywords —Face Recognition, Convolutional Neural Network, Inception, Faster R-CNN

I. INTRODUCTION

Motor vehicle theft is a serious problem because there were more than 100,000 incidents in 2017 [1]. The frequent event of theft is by taking a vehicle in a parking lot [2]. The application of CCTV or live cameras is a mechanism for theft prevention and detection. However, continuous monitoring and vehicle owner recognition are required in the prevention and detection of crime. So even though CCTV was installed, the theft incident still occurred due to negligence in monitoring the parking area. Automotic CCTV monitoring using machine learning methods is solution for suspicious human activity detection on parking lot [3].

Research related to theft prevention in parking lots using detection automation in CCTV has been developed. Najla develops abnormal event recognition in the parking lot for theft preventing [4]. Sayma developed automatic notifications to security for vehicle movements [5]. The use of detection of moving objects can be late to provide theft prevention due to the thief is already in the vehicle. One way to prevent thieves from entering the vehicle is to identify the face of the vehicle owner. The face recognition can be used to check whether the driver is authenticated or not.

The increased interest in the face recognition field has explicitly a relationship with technology that has been available after 30 years of development [6]. Humans have an excellent ability to recognize a person's face regardless of age, lighting conditions, and diverse expressions. Research conducted in this field has a goal to design a facial recognition system that can rival or even exceed the level of human recognition ability, which reaches almost 97.5%. In its application, the methods for facial recognition are deep learning techniques [7].

Deep learning, specifically the Convolutional Neural Network (CNN) method, has achieved excellent results in the field of face recognition. Unlike traditional methods that are made manually, CNN learning features could handle complex intra-personal variations [8]. In face recognition research using a dataset from Labeled Faces in Wild (LFW), CNN has a high degree of accuracy, compared to Joint Bayesian and Local Binary Pattern (LBP) [9].

CNN has a high degree of accuracy using LFW data, but in practice, an image can have several face; therefore, a selective search is performed to determine the location and identification of face. Regional CNN (R-CNN) method is a viable method for a selective search. The results of each selective search were then carried out by CNN [10]. However, the R-CNN method has a weakness, where the computation process runs slowly because of the CNN process is repetitive. This problem is solved by doing CNN on the image first, then determining the location and label of the object. The development of this method is known as Fast R-CNN, and then F-R-CNN Faster. The Faster R-CNN is a development of the Fast R-CNN which adds a proposal region to the CNN method [11]. In the Faster R-CNN, the region of interest is formed using the Region Proposal Network (RPN), while in its predecessor the Fast R-CNN region of interest is formed using external methods such as selective search. The RPN provides more accurate region of interest results, reducing the number of inappropriate regions of interest, thus speeding up the model training process [12]. In its application, Faster R-CNN has various architectures, one of them used in this research is Inception V2. The architecture of Inception V2 is designed to reduce the complexity of CNN [13].

In this research, Faster R-CNN with Inception V2 architecture are used as a face recognition method on CCTV