# Anomaly Detection for Visually Impaired People Using A 360 Degree Wearable Camera

Dong-in Kim and Jangwon Lee
School of Electronics and Information Engineering
Korea Aerospace University
dongpago@kau.kr, leejang@kau.ac.kr

## Abstract

*In recent years, there has been renewed interest in developing Artificial Intelligence (AI)-based assistance systems to help people with visual impairments to overcome their daily difficulties. Such approaches, however, have mostly focused on understanding the contents of an image, and far too little attention has been paid to address physical safety and security concerns of the visually impaired people. Thus, we present a first step towards developing abnormal event detection system for visually impaired people that can observe their surroundings using a 360 degree wearable camera to handle those concerns. In particular, we have two primary aims: (1) to detect any suspicious or abnormal activity in a* $360^\circ$ *panorama panoramic image, (2) to identify the direction of the abnormal event. We introduce a new dataset called* ***AD-360*** *containing 132 videos with a total of 11,037 image frames of abnormal activities such as shoulder surfing and pickpocketing. The videos were collected at three different public places with security and privacy concerns of people with visual impairments in mind. We evaluate our baseline approach with a few slight modifications on the* ***AD-360*** *dataset.*

## 1. Introduction

People with visual impairments face various problems in their daily life. To aid them in those situations, the white cane with a guide dog has been widely used. However, those traditional methods suffer from some serious limitations. For instance, they only can collect limited information surrounding them using the white cane. Moreover, it takes a long time with a high cost to train the guide dogs.

Hence, over the past decade, there has been a dramatic increase in developing a visual aid system that can create new "eyes" for the visually impaired people using wearable devices and Artificial Intelligence (AI) technologies [4]. However, the research to date has tended to focus on understanding some contents/objects in images rather than physical safety and security concerns of the visually impaired people. Therefore, this paper proposes a new methodology for handling physical safety and security concerns of the visually impaired people using a 360 degree wearable camera with AI-based technology. To be specific, we attempts to use 360 degree wearable camera for detecting any suspicious or abnormal activity with its direction in the surroundings of a person with visual impairments. For this purpose, we recorded 132 videos at three different public places (Automated Teller Machine (ATM) booths, cafes, and bus stops) with real-case scenarios based on [1] using the 360 degree wearable camera to build a new dataset called **AD-360**. As a starting point, we design a new model based on the memory-guided architecture [5] with temporal memory blocks and a novel anomaly prediction method to detect abnormal activities with their directions in the 360 degree egocentric videos. We finally evaluate our approach on the **AD-360** dataset and demonstrate that the proposed method is able to detect abnormal activities with directions

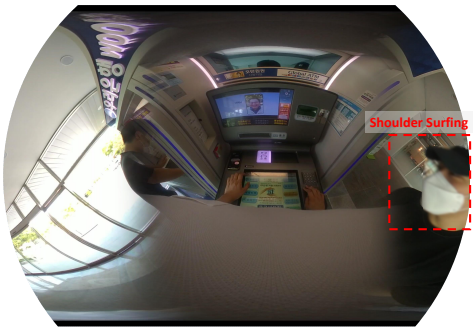


Figure 1. We attempts to detect any suspicious or abnormal activity in surroundings of a visually impaired person. To do this, we collected videos using a 360 degree wearable camera for developing abnormal event detection system. Here is a sample of suspicious activities in our AD-360 dataset: Someone looks over the camera wearer's shoulders to observe the user as he enters his PINs at an ATM booth.

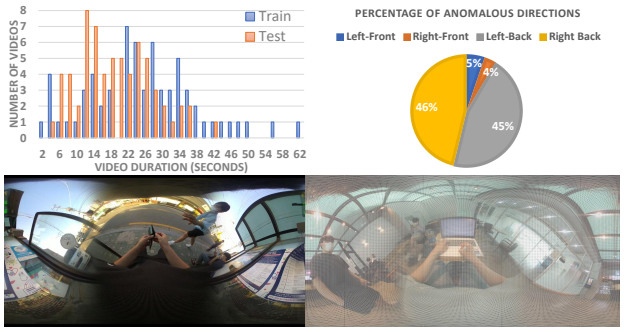


Figure 2. AD-360 dataset statistics (Top) and sample images of anomalous events at a bus station and cafe (Bottom). We overlaid a spherical coordinate on the image frame that is used for our anomalous direction prediction for the cafe image.

in diverse real-life scenarios that the visually impaired people may face in their everyday life.

## 2. Our Approach

### 2.1. The AD-360 Dataset

We introduce a new anomaly detection dataset recorded using a 360° wearable camera dataset (AD-360), the first dataset that contains 360 degree egocentric videos of anomaly activities which visually impaired people often face in their daily life. To construct this new dataset, we collected 132 video clips using a 360° egocentric camera, and then annotated a total of 11,037 image frames to have spatio-temporal labels for abnormal events. The dataset has two types of labels: (1) direction: the direction of anomaly events that should trigger an alert (four directions: Left-Front, Right-Front, Left-Back, and Right Back), and (2) time: when the abnormal events occurred. To build more realistic dataset, we have carefully reviewed the papers that contain the various interviews with visually impaired people about the problems that they face in daily life [1, 3]. We then finally chose a few realistic scenarios that has frequently happened at three different places (ATM booth, cafe, and bus station) based on their interviews for our video recording. Consequently, five subjects wore 360-degree egocentric camera, FITT360 wearable neckband camera [2], to construct the AD-360 dataset based on the selected scenarios at three different places. The videos are recorded at 30 fps at 3840 x 1920 resolution. Fig. 2 shows a few statistics and two sample images of anomalous events at a bus station and cafe.

### 2.2. Detecting Anomalous Events with Directions

We formulate our two tasks as unsupervised learning problems, so that our model does not require any ground truth labels at a training phase. We extend the memory-guided model [5] with temporal memory blocks to capture temporal information in videos, and design the new prediction method to focus on the reconstruction error in a wider area of the image. To identify the directions of anomalous events, we measure structural similarity between the input video frame and the reconstructed frame by prediction using SSIM in the four directions of the camera wearer.

| Method (k = # of input frames) | At Three Different Public Places | | | |
|---|---|---|---|---|
| | ATM Booth | Cafe | Bus Stop | Avg. |
| Memory-Guided Model (K=4) [5] | 62.22 (74.76) | 51.89 (91.14) | 80.50 (56.84) | 64.87 (74.25) |
| Memory-Guided Model (K=8) [5] | 64.49 (73.90) | 49.99 (90.26) | 77.44 (62.14) | 63.97 (75.43) |
| Memory-Guided w/ Temp. Mem. (Ours, K=8) | **68.47 (76.79)** | **63.11 (92.24)** | **82.65 (62.98)** | **71.41 (77.34)** |

Table 1. Evaluation on the AD-360 dataset using AUC(%) and (Accuracy)(%) metrics for anomaly detection and anomalous direction prediction respectively.

## 3. Experiments

Table 1 shows the quantitative results for our two tasks. We measured AUC for anomaly event detection and Accuracy (the total number of correct predictions divided by the total number of anomaly image frames) for anomalous direction identification task.

## 4. Conclusion and Future Work

This paper proposed a novel abnormal event detection task with a new dataset for visually impaired people using a 360 degree wearable camera for the first time. Since we just took the first step in solving this problem, further research should be carried out to explore this new task.

**Acknowledgement:** This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1G1A1010720).

## References

[1] Tousif Ahmed, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia. Addressing physical safety, security, and privacy for people with visual impairments. In *Symposium on Usable Privacy and Security*, 2016. 1, 2

[2] FITT360. https://www.fitt360.us/ (online), 2021. 2

[3] D Gurari, Q Li, Chi Lin, Yinan Zhao, Anhong Guo, Abigale Stangl, and Jeffrey P Bigham. Vizwiz-priv: A dataset for recognizing the presence and purpose of private visual information in images taken by blind people. In *CVPR*, 2019. 2

[4] Danna Gurari, Qing Li, Abigale J Stangl, Anhong Guo, Chi Lin, Kristen Grauman, Jiebo Luo, and Jeffrey P Bigham. Vizwiz grand challenge: Answering visual questions from blind people. In *CVPR*, 2018. 1

[5] Hyunjong Park, Jongyoun Noh, and Bumsub Ham. Learning memory-guided normality for anomaly detection. In *CVPR*, 2020. 1, 2