

Access the code, data, and analysis at [github repo](#)

# Using computer vision and AI techniques to improve crowd safety management

Project report - January 2024

Anton Irvold  
University of Southern Denmark  
[anirv20@student.sdu.dk](mailto:anirv20@student.sdu.dk)

Christoffer Krath  
University of Southern Denmark  
[chkra19@student.sdu.dk](mailto:chkra19@student.sdu.dk)

## Administrative Info

### Collaborators from Event Safety

- Sofie Dahl, [sofie@eventsafety.dk](mailto:sofie@eventsafety.dk)
- Christian Sejlund, [christian@eventsafety.dk](mailto:christian@eventsafety.dk)

### SDU supervisor

- Sune Lundø Sørensen, [slso@mmmi.sdu.dk](mailto:slso@mmmi.sdu.dk)



## Context and Background

As seen from Figure 1 accidents, fatalities and injuries at festivals and concerts worldwide are on the rise, with just below 5,000 worldwide deaths in the decade of 2010-2019. This can be attributed to several causes, according to Raineri (2004): First of all, an increase in the popularity of outdoor music festivals resulting in more and larger festivals with large crowds. Also, high-risk behavior among crowd attendants and the performing artist's music, behavior, and stage show affects the crowd's safety. Cultural influence has also played a large part in the safety of outdoor musical events. This can be behavior such as crowd surfing, moshing, stage diving and the like. Sudden panic in a crowded place can thus affect the safety of the crowd.

It is worth mentioning that these figures and descriptions of situations are from media outlets. As such, it's not the full story. While these statistics might make it seem that general safety and incidents at festivals and the like have gotten worse, it is the opposite, It is generally incredibly safe to attend these events, but using newer technologies we will attempt to enhance it further.

At the same time, festivals in Denmark and the rest of the world have had a focus on mitigating the safety risk in large crowds. This can be seen by the establishment of the [Event Safety Foundation](#) in 2015. This is a joint venture between the Skanderborg Festival Group (Smukfest) and Muskelsvindfonden (Grøn Koncert). This foundation has the purpose of securing the safety of the crowds at these two large Danish festivals, as well as many other events in Denmark. They do this through knowledge sharing, professionals and trained volunteers, courses and counseling. This is the company that this project will be done in collaboration with.

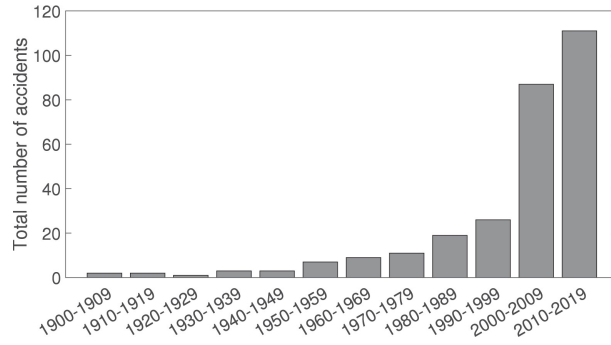
To have the best possible outcome for this project, Event Safety invited us to Grøn Koncert and Smukfest in July and August to see how they work and gather security video footage of real crowds at festivals to use as training data. We will also use Event Safety for their practical and theoretical knowledge of crowd safety to use in the system, user feedback and user testing.

## Problem Statement

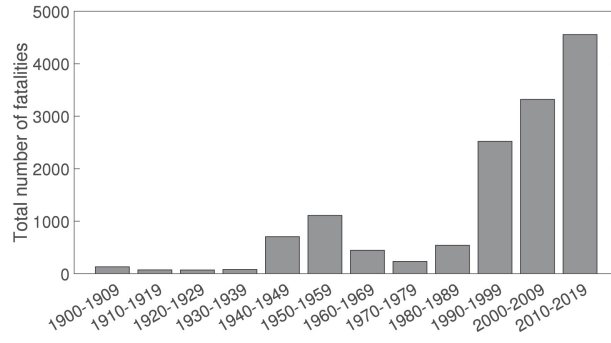
Can computer vision software and AI techniques be leveraged to improve crowd overview for security guards, by receiving video feed from large crowds, and ultimately improve crowd safety?

## Solution Proposal

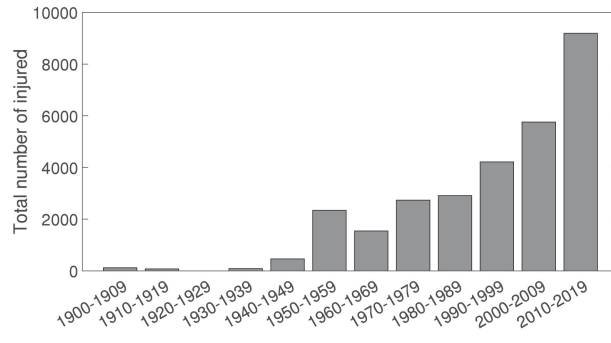
Using cameras mounted around a hotspot of a crowd or on a stationary drone, we aim to develop a platform where an AI model receives live footage from these cameras, uses the model to segment the crowd, and learns what a dangerous situation might look like. This could be one or more of the following (Raineri, 2004):



(a) Accidents by decade



(b) Fatalities by decade



(c) People injured by decade

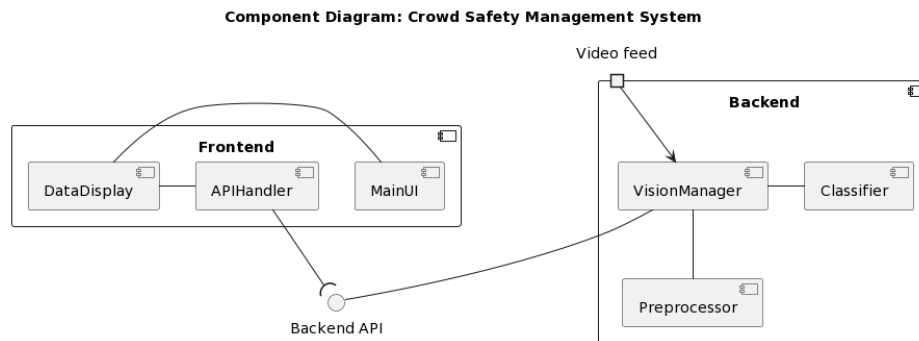
**Figure 1:** Fatalities in crowds graph (Figure from Feliciani et al., 2023)

- Several people moving into the crowd from a specific direction create a dangerous pressure point
- More than a specified amount of people in a marked area resulting in unsafe conditions (ie.  $>6$  people per square meter)
- Omnidirectional or directional movement in the crowd resulting in a dangerous situation
- An area in the crowd suddenly being void of people, perhaps hinting at a mosh pit or an emergency

These are some of the situations this project aims to systematically detect and alert professionals by providing meaningful feedback.

The system would consist of the following:

1. A backend that receives the data from the cameras developed in either Java or Python, with an implementation of a machine learning model to handle the video feed according to the bullet list above and exposing this data through an API
2. A simple frontend or GUI (Graphical User Interface) to display this information in a meaningful way. This could be through a heat map, a numerical estimate of a risk factor, or other visual output.



**Figure 2:** Component diagram of the system

An initial component diagram that maps this described system can be seen in Figure 2.

## Timeline

Date	Activity
July-August	Gathering data at attended festivals in collaboration with Event Safety

Date	Activity
August	Drafting the project description
31-08-2023	Delivering the project description
09-09-2023	Final discussion with Event Safety regarding the proposed solutions, requirements and the project going forward
September	In-depth analysis of use cases (with Event Safety) In-depth analysis of required technologies and methods used in academic literature. Design of the system
25-09-2023	Progress update with Event Safety regarding the design of the system
October	Implementation of proof of concept system (primarily backend)
26-10-2023	Progress update with Event Safety and small-scale user testing
31-10-2023	Attending Crowd Safety Course hosted by Event Safety in Copenhagen
November	Implementation of proof of concept system
December	Final implementation of the system Documentation of the system User testing and evaluation

# Software Requirement Specification

System description

Theory and methods

Project planning

Acronyms and definitions

Acronyms	Definition
CSM	Crowd Safety Management System
AI	Artificial Intelligence
ML	Machine Learning
CCTV	Closed Circuit Television
Crowd Surges	A dangerous crowd situation where your movement is controlled by the crowd and not your own actions.
Mosh pit	A crowd situation where people vacate an area in the crowd followed by violent or aggressive dancing. Sizes of Mosh pits vary greatly.
Technical user	A developer or technician that can implement the CSM in the technical setup of a concert or venue
Non-technical user	A security guard or someone viewing crowd footage through the CSM without any technical or software-related background
Choke point	Point with high crowd density and crowd influx

## System specification

### Functional Requirements

ID	MoSCoW	Requirement
1	M	The system <b>must</b> be able to count the amount of people in the crowd
2	M	The system <b>must</b> be able to segment the crowd into virtual sections of equal physical size
3	M	The system <b>must</b> be able to create a heatmap of the crowd density
4	S	The system <b>should</b> be able to detect the movement of the virtual crowd sections
5	S	The system <b>should</b> be able to distinguish different types of movement
6	S	The system <b>should</b> be able to detect choke points in the crowd movements
7	S	The system <b>should</b> be able to generate a summarizing report of the concert with statistics of (crowd density, crowd count, risk factors, choke points, and other) relevant safety indicators after the event
8	C	The system <b>could</b> be able to estimate a numerical risk factor based on available factors
9	C	The system <b>could</b> be able to generate a live, or delayed, video overlaid user interface
10	W	The system <b>will not</b> be able to identify dangerous situations that might call for cautionary actions such as crowd surge, mosh pits, falls, blocked exits, etc.
11	W	The system <b>will not</b> be able to use data gathered elsewhere at a concert such as alcohol sales, average crowd age, sound, artists, etc. to give a more precise crowd profile and thus risk factor

### Non-functional requirements

ID	MoSCoW	Requirement
1	M	The system <b>must</b> protect the privacy of the CCTV video feed
2	S	The system <b>should</b> have a user-friendly interface that is easy to manage for both technical and non-technical users
3	S	The system <b>should</b> have adequate documentation / technical specification for technical users
4	S	The system <b>should</b> have adequate user manuals for non-technical users
5	C	The system <b>could</b> have high reliability that is not based on the visual circumstances and environment (e.g. sunlight, stage light, audience flashlights, and other visual effects)
6	C	The system <b>could</b> be scalable to simultaneous interoperability between multiple cameras
7	C	The system <b>could</b> integrate with existing CCTV software systems at venues



**Detailed Requirements**  
**Risks**

## Analysis

### misc

Use Crowdcouting and NOT objectcounting models. Object counting is better for segmenting visually clear definitions and not individual members of a crowd

### Regarding SAM model for crowd counting:

“Although the Segment Anything model (SAM) has shown impressive performance in many scenarios, it currently lags behind state-of-the-art few-shot counting methods, especially for small and congested objects. We believe that this is due to two main reasons. Firstly, SAM tends to segment congested objects of the same category with a single mask. Secondly, SAM is trained with masks that lack semantic class annotations, which could hinder its ability to differentiate between different objects. Nevertheless, further exploration of adapting SAM to the object counting task is still worth studying.” (Ma et al., 2023)

### Completely Self-Supervised Crowd Counting via Distribution Matching

Unsupervised Learning, Crowd Counting, Deep Learning

Possible to use a model not trained on relevant data but still able to produce meaningful results. Some are even better than other models. (Sam et al., 2020)

### Possible solution using google colab

Setup an environment implementing SASNet. Eliminates use of anaconda but perhaps at the cost of a paid subscription.

It was considered to use a CNN model called CRSNet. However, research suggests that CRSnet might have problems with scale variations, which is a common problem in crowd counting. Because of this, SASNet is quite possibly a much better solution. (**sasnetCrowd**) (**sasnet**)

## Design

## Implementation

## **Validation and Verification**

## Discussion

## Conclusion

## Perspective



## References

- Feliciani, C., Corbetta, A., Haghani, M., & Nishinari, K. (2023). Trends in crowd accidents based on an analysis of press reports. *Safety Science*, 164, 106174. <https://doi.org/https://doi.org/10.1016/j.ssci.2023.106174>
- Ma, Z., Hong, X., & Shangguan, Q. (2023). Can sam count anything? an empirical study on sam counting.
- Raineri, A. (2004). The causes and prevention of serious crowd injury and fatalities at outdoor music festivals. <https://doi.org/10.13140/2.1.3036.0005>
- Sam, D. B., Agarwalla, A., Joseph, J., Sindagi, V. A., Babu, R. V., & Patel, V. M. (2020). Completely self-supervised crowd counting via distribution matching.