Crowd counting and monitoring videos for surveillance video.

ABSTRACT:

In public venues, crowd size is a key indicator of crowd safety and stability. Monitoring the people number and crowd density levels are important. In this , we present a framework that will enable real-time crowd counting and spatial-temporal analysis for the crowd of the monitoring region. Firstly, we will obtain crowd counting models for each camera by statistics regression methods using sample data. Secondly, we will integrate video surveillance system and geographic information system (GIS) for capturing, managing, analyzing and displaying all forms of geographically referenced camera information, such as location, monitor area, and real-time crowd counting data, etc. And then, we combine image processing with crowd counting models to estimate people number and crowd density of monitoring areas. Finally, we implement a system for real-time crowd counting based on video surveillance system and GIS. Tracking is based on a probabilistic approach in which the appearance and probability of occlusions are computed for the current camera and warped in the next camera's view by positioning the cameras to disambiguate the occlusions. We can acquire real-time data of people number and crowd density levels for each camera, and display them by the way of map and curves using Kibana. Also, we can retrieve history data and analyze them by spatial analysis tools. The experiment shows that this system can provide early warning information and scientific basis for safety and security decision making.

INTRODUCTION:

With the accelerated process of urbanization and socio-economic development, people gather in public more and more, which results in many crowded public emergencies. For reducing the number of crowd accidents, estimating the people number and crowd density should be very pivotal. As the crowd counting and crowd density are essential descriptors of all the status of crowd, it is very useful for us to learn the distribution of crowd and find out the tendency of abnormal behaviors, if we know more about the people number and density of the crowd. So crowd counting and crowd density are very useful information for security departments.

Detecting and counting people in surveillance systems is a challenging problem. In many years, there were many works dealing with estimating the people number and crowd density for crowds. But most of them were related to algorithm research for a single camera. The goal of the crowd density detection is the basis for security sector's decision-making. During the emergency evacuation, potential factors are overcrowding and crushing caused by, for example, human stampede behaviors and structural problems of pedestrianfacilities. In order to design effective and safe pedestrian facilities and egress routes, it is important to spatially and temporally understand the pedestrian behaviors and the egress efficiency.

GIS is a system to capture, store, manipulate, analyze, manage, and present all types of geographically referenced data. In this , we design and implement a system that integrated GIS with video surveillance system. It can capture and analyze real-time crowd counting data in spatial-temporal pattern and give security sector useful information to help them
make scientific decisions.

In some scenarios, such as public rallies and sports events, the number or density of participating people is an essential piece of information for future event planning and space design. Good methods of crowd counting can also be extended to other domains, for instance, counting cells or bacteria from microscopic images, animal crowd estimates in wildlife sanctuaries, or estimating the number of vehicles at transportation hubs or traffic jams, etc.

PROPOSED SYSTEM:

This system is composed of two components :

1.Crowd counting:

This kind of crowd counting methods are to output a density map of the crowd and then obtain the head count by integration. The essential work is accurate estimation of the crowd density distribution. The main objective is to learn a mapping between an image I and corresponding density map D,for a m pixels by n pixels image. For a dataset, assuming that the position of each pedestrian is labeled by dot at the center of every head, the density ground truth D is obtained by creating a sum of Gaussian kernel at the center of every head. The number of crowd is calculated by integrating the estimated density map.

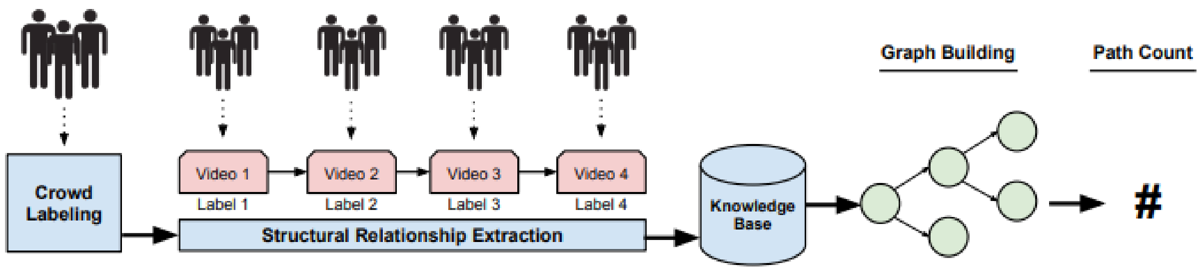
2.Video surveillance:

A video surveillance system is the use of video cameras to transmit signals to a specific place.To design a system framework of real-time monitoring crowd density combined video surveillance system with GIS.

TECHNOLOGY STACK:

1. Python based Computer Vision and Deep Learning libraries
2. OpenCV
3. Keras
4. TensorFlow
5. YOLO
6. CNN
7. Machine Learning

DEPENDENCIES:



First, video is decomposed by the crowd into a set of videos, each containing a single mid-level action, and then an action label is found. Next, Architect recruits crowd workers for each labeled video segment (action). These workers are asked a series of randomly ordered questions pertaining to the required ordering of pairs of tasks. Finally, the set of pair wise restrictions are used to create a dependency graph. This information can be used to help the system recognize familiar actions performed in orders it has never seen before.

METHODOLOGY

Step 1: Data collection and dataset preparation

This will involve collection of publicly available video dataset from available sources. There are many publicly available datasets like UCSD pedestrian dataset, Grand Central railway station dataset etc.

Step 2: Developing a CNN based Crowd counting and monitoring model

To count the people from crowd video, a CNN based model will be developed. The model will have convolution layer, RELU and max pooling layer, fully convolutional layer and Softmax activation function. Firstly, individual frames will be extracted from the video and passed as input to the convolution layer. A convolution of appropriate size (M x N) will be applied on input image to extract the features and create a feature map. The model will be trained by some sample training data and then it will be tested by some test data. The feature extraction will be performed automatically by the model. When training and testing part is over, new input video will be applied and output of the model will be matched with ground truth values.

Depending on the accuracy of the result more number of layers will be added or model will be trained against more number of epochs. Popular pretrained CNN feature extraction models such VGG16, ResNet or FCN (Fully Convolutional Network) as shown in fig 2 will be exploited for this task.

Step 3: Training and experimentation on datasets

The Crowd counting and monitoring model will be trained both on the large-scale datasets such as UCSD pedestrian dataset, Grand Central railway station dataset and live recorded videos populated based on CCTV recording camera as part of this project. The part of the dataset will be divided into training data and rest of the data will be used as testing data.