ABSTRACT:

PROBLEM STATEMENT:

Visual pollution needs to be defined, formalized, measured, and evaluated from a variety of perspectives because it is a relatively new problem in comparison to other types of environmental contamination. Through the use of 21st-century technology to environmental management applications, this competition seeks to build a new sector of automated visual pollution categorization. We anticipate competitors to imitate the human learning process in the area of picture identification for the categorization of visual contaminants by training and testing methods for convolutional neural networks. The creation of a "visual pollution score/index" for urban areas would also benefit from this since it might result in a new "metric" or "indicator" for the field of urban environmental management. We will create and improve algorithms in this competition using a sizable dataset. The raw sensor camera inputs as seen by a fleet of many vehicles in a constrained geographic area in KSA are included in this collection.

SOLUTION TO THE STATEMENT:

2 Related work

The foundation for the system examined in this research is briefly described in this section. As previously mentioned, the primary innovation of our system is the integration of raw picture segmentation with optical flow segmentation. Additionally, these procedures have to ideally be performed in real time. As a result, it is important to put into perspective a number of the application areas involved, including the parallelization of calculation techniques and the segmentation of findings.

Every algorithm has various benefits and drawbacks, but the majority of them are mostly hampered by high computational and memory costs. Some of them attempt to do so by compromising the precision of the findings; in other words, they strike a compromise between the algorithm's cost and degree of accuracy.

Many studies have been conducted in the area throughout the years. It has been continually enhanced, sometimes by focusing on the algorithm itself, other times by fusing two of them, and other times by fusing with other methods.

CLASIFICATION:

The Mark 1 Perceptron machine, which was used for picture categorization, was created in the 1950s using the perceptron method, making it the first computer vision job to be addressed by machine learning.

This technique proved effective for structured data issues, but it was limited to simple tasks like categorizing various geometric forms. A few decades later, the SVM algorithm was developed, which could deal with high-dimensional data, such tiny picture datasets, with the fewest possible samples.

The introduction of convolutional neural networks (CNNs) by Yann Lecun in his model LeNet, which was shown to be superior to other vision-based ML techniques in 2012, when AlexNet was the first CNN-based model to win the renowned ImageNet competition, is what has really revolutionized computer vision.

Localization and Object Detection

A computer vision issue called object localization and detection requires the algorithm to determine the positions of one or more target objects in an image and produce bounding boxes for each one that appears in the image or video frame.

This activity is heavily utilized in a variety of fields, including autonomous vehicles, robots, augmented reality, and medical applications.

Problem Definition

Depending on the architecture being utilized, this challenge can be expressed in a variety of ways. However, typically, the network should use a sigmoid or softmax activation function to produce a class for each target item in the picture.

Additionally, the network produces four variables that can be (x, y, w, h) and indicate the bounding box's center or top right corner, respectively, in order to locate the item.

The bounding box prediction is a regression issue. The output is created for different areas of the picture either via CNN implementation or via a sliding window, which means that the majority of algorithms need producing another variable indicating whether an item exists in the specified region or not.

R-CNN:

The R-CNN strategy has been adopted by several architectures based on region recommendations. As it combines feature extraction and classification into one CNN, Fast R-CNN performs and runs faster than R-CNN. As a result, the network receives an image and many ROIs as input and generates predictions for the class and bounding box for each ROI.

By separating ROIs from the network, faster R-CNN went one step further, increasing both accuracy and speed. This is because the network is given more latitude to find a solution. The network has complete control over ROIs and feature extraction since the weights may be adjusted across the network via end-to-end training.

FUTURE SCOPE:

Rapid population growth has primarily raised demand for and use of motorized vehicles everywhere. Road accidents have significantly grown in frequency over the past ten years as a result of this rise in motor vehicle use. Despite the safety precautions put in place for the public and private transportation system, the number of injuries, disability, and fatalities caused by deadly road accidents has been rising every year. The actual causes of accidents across the world include traffic congestion, distracted driving, street racing, impaired driving by drugs or alcohol, poor traffic light or vehicle design, tailgating, running stop signs and red lights, making incorrect turns, and driving in the wrong way. Road safety is monitored by a variety of sophisticated monitoring systems, but there is still considerable room for improvement. Although there are numerous cutting-edge monitoring systems in place for road safety, there is still considerable space for improvement in terms of accident prevention. Accidents can be avoided by using the advanced car tracking and traffic surveillance systems already in place. Real-time observations are challenging, however, due to the massive volume of continuous monitoring data. Using cutting-edge real-time technology might aid in accident detection and prevention given current advances in information and computer science.

This technology that aims to replicate how the human visual system functions is known as computer vision. Real-time digital picture data collection and analysis are performed on a variety of surveillance systems. The systems simultaneously identify and notify any occurrences, such as speeding, irresponsible driving, accidents, etc. Some computer vision-based techniques using cutting-edge deep learning methods that may be utilized for real-time accident detection and prevention procedures include image classification, object recognition, object tracking, semantic segmentation, and instance segmentation.

Similar to this, various irregularities in the movement of cars may be found using neural networks and historical data, which can also be used to prevent accidents. The use of deep learning techniques for visual identification has recently made great strides, and this has had a big influence on advanced computer vision research. Furthermore, real-time traffic monitoring for the purpose of preventing and detecting accidents would greatly benefit from the use of computer vision. The objective of this special issue on "Real-time

Computer Vision for Accident Prevention and Detection" is to provide the most recent advances in real-time computer vision and image processing for the purpose of accident prevention and detection.

OUTPUT :

The LOP
KINEMASTER
pothole: 86%
pothole: 99%
0
100
200
300
400
0
100
200
300
400
500
600
700
800