SURVEILLANCE

Prof.S.G.Lade, Department of Computer Engineering, Vishwakarma Institute of Technology

Prof.D.R.Deshpande, Department of Information Technology, Vishwakarma Institute of Technology

Amartya Mathew (C-50)
Department of Computer Engineering Vishwakarma Institute of Technology, Pune

Nikhil Kulkarni (H-24)
Raghav Sadany (H-09)
Department of Information Technology Vishwakarma Institute of Technology, Pune.

Abstract – A prototype of a novel Surveillance system which involves mounting cameras on an unmanned all-terrain vehicle (ATV). The main aim is to detect movements of objects while the cameras are in motion and try to classify what has caused the motion using image processing. This will be a good tool for surveillance in areas where having human personal is risky or not possible.

Index Terms -Motion Detection, Object Recognition, Communication, Android Interface, Convolutional Neural Network, Optical Flow.

I. INTRODUCTION

Country borders and war areas yield inevitable injuries and fatalities. Many of the recipients are recognized for their outstanding devotion and commitment to the country. To assist the armed forces in regions where surveillance is too dangerous or difficult we propose a prototype of a novel Surveillance system which involves mounting cameras on an unmanned all-terrain vehicle(ATV) that takes video inputs and sends them to the system. The system detects motion from a streamed video file and analyses the frames in real-time, required that various aspects to be taken into consideration. Some of these aspects are in contrast created by weather outside, flying machines that may fly over a restricted zone and strong winds that can affect stationary objects to activate the system.

So instead of a person performing surveillance we can mount a camera on any unmanned vehicle (ATV) and by image processing data required can be obtained. The System functioning has four major steps:

1 Optical Flow(Motion Detection)
2 Convolutional Neural Network(Image Classification)
3 Firebase(Storage/Accessing Data for Remaining Operations)
4 Android Application(For the User to have access to the data with a better interface)

II. THEORY

A. Optical Flow

Optical flow is the pattern that can be detected due to the apparent motion of image object between two consecutive frames caused by the movement of object or camera.
It is a 2D vector field where each vector is a displacement vector which has the info of the movements of points from one frame to the next as seen in figure 1.

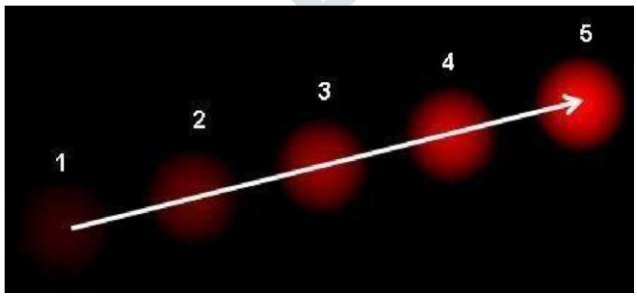

Figure.1 Optical flow frames. Source: image1

Figure.1 shows a ball moving in 5 consecutive frames. The arrow shows its displacement vector. Optical flow has many applications in areas like:

1 Structure from Motion
2 Video Compression
3 Video Stabilization

Optical flow works on several assumptions:

1. The pixel intensities of an object do not change between consecutive frames.
2. The neighboring pixels share a similar motion.

B. Convolutional Neural Network(CNN)

Convolutional Neural Networks (ConvNets or CNNs) are a subset of Neural Networks that have proven very effective in areas where the data can be represent as a matrix. Convnets have been very successful in identifying objects in images in some cases even exceeding human efficiency. Convnets success in these tasks make them ideal for applications like self-driving cars and security surveillance. The figure below shows an example of ConvNets being used for recognizing everyday objects, humans and animals. Lately convnets have been used for audio tasks and Natural Language Processing tasks (such as sentence classification) as well.

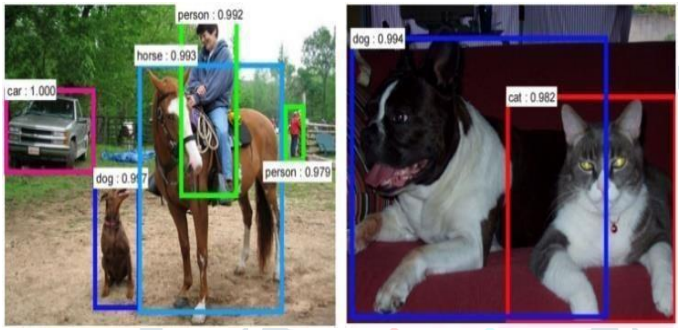

Figure .2 ConvNet recognized images. Source : image2

The four main mathematical operations in the ConvNet are:
1) Convolution
2) Non Linearity (ReLU)
3) Pooling
4) Classification using Fully Connected Layer

C. Firebase

Firebase is a platform provided by Google which provides storage, database facilities for Android applications. Firebase storage facility allows developers to securely upload and download files for different purposes, regardless of network quality. The storage provided by firebase is backed up by Google Cloud Storage. The Data obtained from the motion detection algorithm is used by the Convolutional Neural Network to detect the cause of the motion. Thus to send this processed data to the user via an Android application we have used Firebase to store and retrieve the data.

D. **Android**

Android Application is the last stage of our system where we display the final processed images from the motion detection and the CNN based classification model to the user. Thus the users gets to view all the images where a motion is detected as well as the cause of the motion on an Android Interface.

III. ALGORITHM

1. 221 equidistant pixels (17*13) in the video frame are set to be considered for optical flow each being 20 pixels apart.
2. Every successive frame in the video gives the displacement of those 221 points in the new image using optical flow.
3. We add the values of the average of all the displacement into a queue of size 10.
4. Using this the threshold after every 10 frames is automatically refreshed.

final_threshold = final_thresh/10

5. Now, a different queue is created to add all the displacements of 221 points of every frame to check their displacement with respect to the automated threshold value received before threshP = LP.get()

6. We set the minimum and maximum number of points required for detection.

7. We do this to get accurate results and reduce the unwanted noise due to various factors.

8. The detected images are now sent to the firebase.

IV WORKFLOW

The calculation of threshold is done automatically as it takes an average of the first ten frames and applies the value to the next ten frames and the process goes on. As a result the threshold refreshes after every 10 frames. Thus the model can work on variable speed of the ATV and still give accurate results.

The calculation to find out the optical flow between the equidistant pixels:
Consider a pixel I(x, y, t) in first frame .As the time is negligible between two consecutive frames, in the next frame that point moves by distance (dx, dy) taken after time dt. So since the intensities of the group of pixels don't change considerably, we can say,

$I(x,y,t)=I(x+dx,y+dy,t+dt)$

Applying Taylor series approximation of right-hand side, remove common terms and divide by dt to get the following equation:

$fxu+fyv+ft=0$ where,
$fx=\partial f\partial x;fy=\partial f\partial y$

$u=dxdt;v=dydt$ ……(1)
Above equation (1) is called Optical Flow equation. We can find fx and fy, they are image gradients.

Similarly ft is the gradient along time. As (u, v) is unknown, this one equation with two unknown variables is unsolvable. As a result we always take kernel of group of pixels while performing Optical flow.

A. Motion Detection

The Python program for Motion Detection is responsible for the data extracted. When any object is in motion faster or slower than the set threshold of the code then that particular frame is extracted and stored in the form of an image. This image is transferred on the online storage platform used Firebase.

The work flow of the code is such that, it compares two frames and plots fixed points on each frame. The change in positioning of the points helps us in calculating in the displacement of the points.

B. Image Classification

The images uploaded on Firebase are accessed for their classification. By using CNN we first train the machine with a Dataset using Back propagation for recognizing and classifying them. This trained code is then used in classifying the images of firebase .If the images are of concern to the user then it is labelled and uploaded on Firebase according to the categories for example human, animal, machines etc.[5][6]

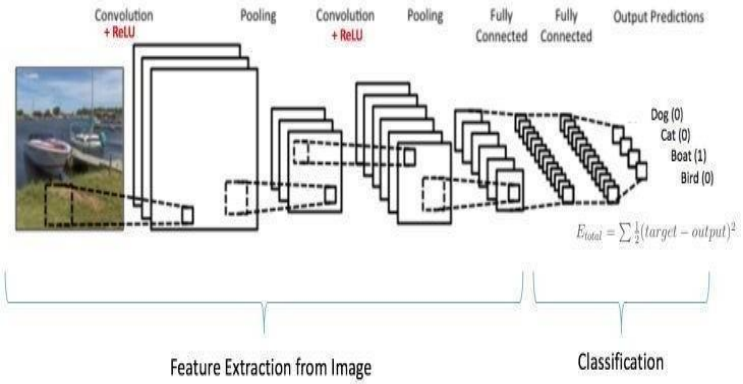

Fig.3 Object classification Flow chart Source: image3

The overall training process for making the convnet is done in the following way step 1: All the filters and

weights are randomly initialized.

Step 2: The network takes a training image as an input and goes through the forward propagation step that is the Convolutional, ReLU, Pooling and then the fully connected layers. All this leads to the output of probabilities of each class being in the image.

The probabilities will be random due to the random nature of the filters and weights.

Step 3: The error of each class is calculated and the summation is taken. The total error is calculated.
Total Error = $\sum$ ½ (target probability – output probability) ²

Step 4: Backpropagation is used to calculate the gradients with respect to the all the weights.
Gradient descent is used to update the values of the weights so as to minimize the output error.

1) The weights are now adjusted.

2) So if the same image is given again the network will be able to classify is closer to the actual expected output.

3) This means that the network has learnt to classify the particular image.

4) Only the filters and weights of the network get updated the other parameters that are decided before the training begins stay the same.

Step 5: Run steps 2-4 with all images of the training set. The above steps of the training mean that all the weights and parameters of the convnets have been optimized to correctly classify images from the training set. When a new image that the convnet has not seen is presented the network will propagate the image through it layer and then output probabilities. If the training set is large enough the network will be able to generalize well enough for the new images.

C. Android and Firebase interface

After the classification is done the images are again posted on Firebase. The Android Application directly gets access to the images posted on firebase with its simple and user friendly interface. The Application is developed using Android Studio.

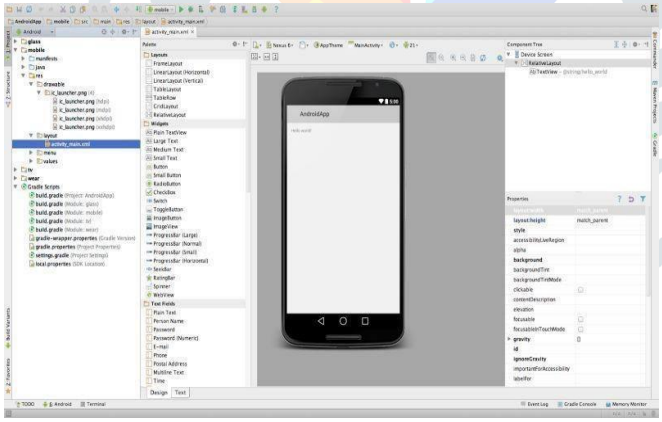

Figure 4 Android interface Source: Internet

Figure 5 Output of Detected Motion using Motion detection algorithm

V FUTURE SCOPE

1. The code can be run on Raspberry Pi module 3. It has an inbuilt WIFI module on it. A camera can be also mounted on it and a robot vehicle can be controlled using Rpi. The only problem is that there is a hardware restriction. The RAM of Rpi is not sufficient enough to run the Program in a sufficient way. It is very slow as compared to that on Computer.
2. A vast area as Computer Vision can be used as data providing source in many applications. It has a huge scope in Surveillance as well cause of automating the whole monitoring system.
3. Better communication techniques can be used like GSM Module for access to the data from system in remote areas.
4. GPS can be used for the accurate coordinates for locating the threats faster and efficiently using the available resources.
5. Night vision capabilities can be added by using IR sensors.

VI CONCLUSION

Surveillance has been performed without the presence of any human. This concept can be used in many areas to detect movement where continuous monitoring is required. Currently the software phase of the algorithm is complete whereas the hardware implementation is in development phase.

VII ACKNOWLEDGEMENT

We would like to thank our project mentors Prof.D.R.Deshpande, Department of Information Technology, Vishwakarma Institute of Technology, Prof.S.G.Lade, Department of Computer Engineering, Vishwakarma Institute of Technology for their invaluable guidance and kind cooperation.

REFERENCES

[1] J. Khurshid, “Military robots – a glimpse from today and tomorrow” 2004.
[2] A.Nayyar, “Smart Surveillance Robot for Real-Time Monitoring and Control System in Environment and Industrial Applications” 2018.
[3] S. Joshi , “Surveillance robot for military application,” 2016.
[4] Widodo Budiharto ”Intelligent Surveillance Robot with Obstacle Avoidance Capabilities Using Neural Network,” 2015.
[5] M.Karthikeyan, “Intelligent Exploration and Surveillance Robot In Defence Environment” 2014.
[6] David J. Fleet, Yair Weiss, “Optical Flow Estimation”
[7] S.Nalawade,”Robots For Surveillance In Miitary Applications” 2014.