Two-stream Convolutional Neural Network for Natural Gas Leak Quantification

Jingfan Wang (jingfan)
Department of Energy Resources Engineering

Motivation and Problem Statement

Natural gas leaks waste money, reduce energy availability, and result in both local air quality and global climate impacts. The climate impacts of leaked gas are particularly important due to the high global warming potential of methane (a key component of natural gas).

We are performing an interdisciplinary project that expands upon EPA-approved IR imaging and will harness the potential for deep learning advances to allow for rapid automatic classification of methane leaks. This will solve the problem of labor-intensive optical imaging methods that lack the intelligence of knowing how big a leak is. Long-term RNN and two-stream CNN are two mature techniques for video classification. Applying CNN in quantifying the leak size is a method, which any researchers haven't thought about. I will try to solve the problem of estimating the leak size by the innovative method, two-stream CNN model including spatial and temporal streams.

In this project, we will work to solve two issues:

1. How to develop a two-stream CNN model for estimating the size of leaks?
2. How is the performance of the model? Why is it better than other methods, like using our own eyes?
3. What kind of leak is usually misclassified? Why does it happen?

Datasets and Features

Data collection and datasets

- On campus, I performed controlled release study of natural gas leak, which means when taking the video of natural gas leak, the true label of the leak size is known.
- Totally, I finished collecting 5 sets of videos. In one set of collection, the camera was fixed in the same distance and angle.
- 6 levels of leak were collected (Class 1: 1L/min, Class 2: 2L/min, Class 3: 3L/min, Class 4: 4L/min, Class 5: 5L/min, Class 6: 0L/min).

5 sets of videos * 6 classes of leak * 1495 frames per video = 44850 frames

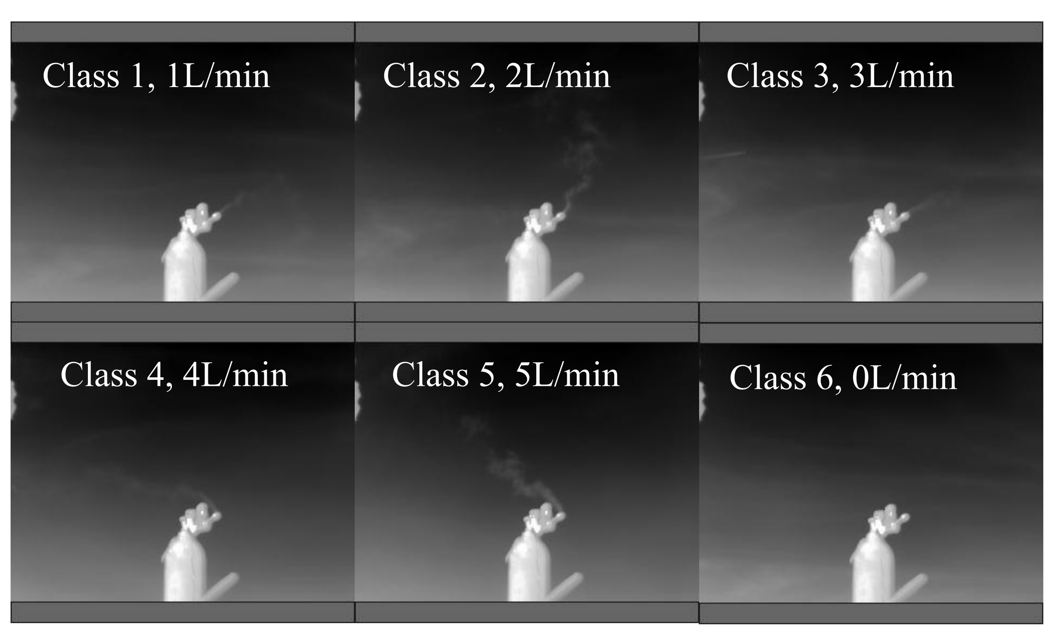


Fig 1. Snapshots of six levels of videos.

Fig 2. Snapshots of three frames in the same video of Class 3.

Input for temporal stream – X and Y Velocity Field from Dense optical flow

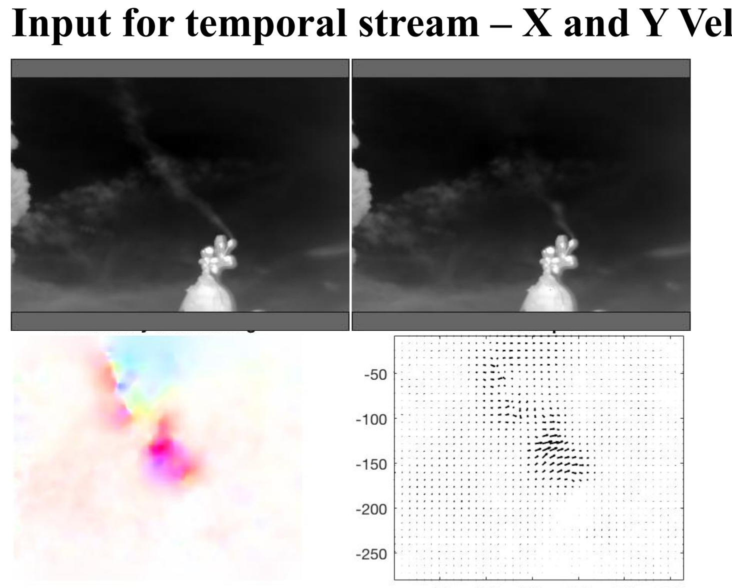


Fig 3. Calculation of dense optical flow

Optical flow: Conservation of intensity of one pixel (x,y,t)

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

Taylor series approximation of right-hand side

$$f_x u + f_y v + f_t = 0$$

Dense optical flow: Apply polynomial expansion to approximate some neighborhood of each pixel with a polynomial

$$f(\mathbf{x}) \sim \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c.$$

Training set and testing set

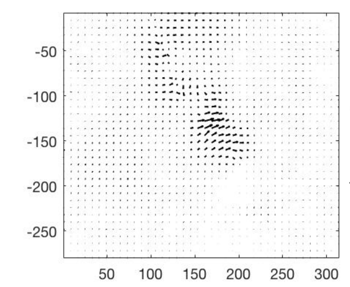
Partition each video into five parts. Randomly choose four of them as training set, one of them as testing set

One video sequence	Training	Training	Testing	Training	Training

Two-Stream Architecture and Results

Two-stream architecture

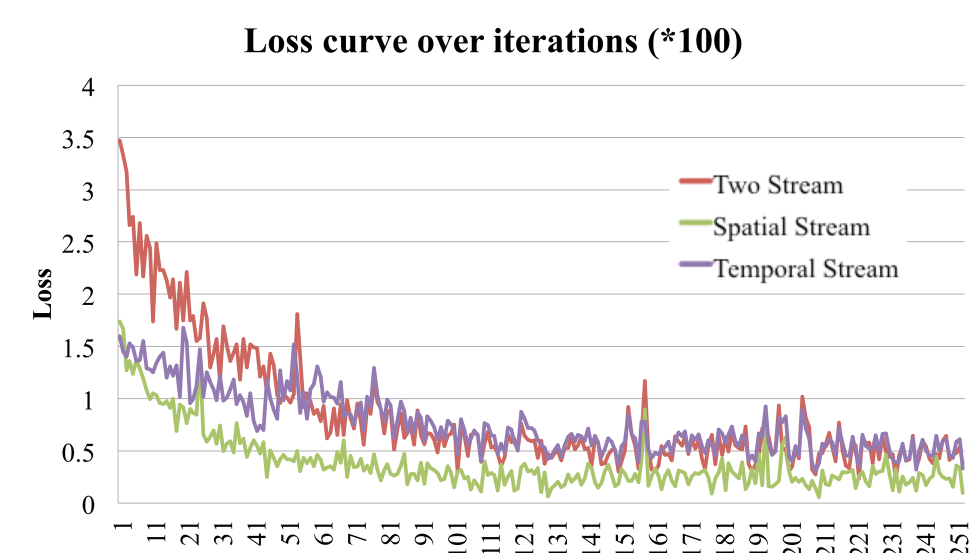
Stream	Conv1	Conv2	Conv3	Conv4	Conv5	Full5	Full7	Softmax
Spatial Stream	3*3*48 BN ReLU maxpool 2*2 stride 2	3*3*96 stride 2 BN ReLU	3*3*256 stride 2 ReLU	3*3*512 stride 2 ReLU	3*3*512 ReLU maxpool 2*2 stride 2	4096 dropout	2046 dropout	
Temporal Stream	3*3*48 BN ReLU maxpool 2*2 stride 2	3*3*96 stride 2 BN ReLU	3*3*256 stride 2 ReLU	3*3*512 stride 2 ReLU	3*3*512 ReLU maxpool 2*2 stride 2	4096 dropout	2046 dropout	



Class score fusion

Fig 4. Two-stream CNN Architecture

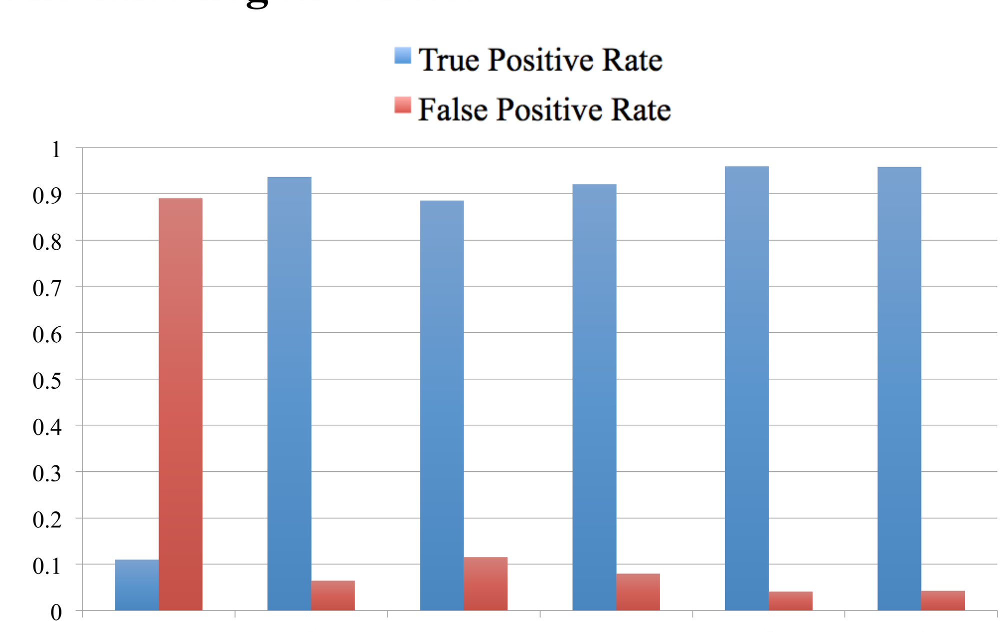
Loss curve and classification results



Architecture	Accuracy: Training	Accuracy: Testing
Spatial Stream	68.30%	65.20%
Spatial Stream on the previous datasets (25 videos taken in 25 angles)	64.90%	61.10%
Temporal Stream	56.50%	52.30%
Two Stream	80.70%	77.60%

Confusion matrix and results of true-positive and false negative rates

True Labels	Prediction 1	2	3	4	5	6
1	164	417	261	213	211	229
2	33	1399	19	19	14	11
3	15	149	1323	2	2	4
4	21	23	67	1376	4	4
5	7	7	7	40	1434	0
6 (No Leak)	13	15	10	18	7	1432



Examples for misclassification

Fig 5. Misclassification on spatial stream

All three images are from Class 3. Due to the wind, sometimes the plume is hidden behind the cylinder, then the prediction would be smaller; sometimes the plume is covering more area, then the prediction would be larger.

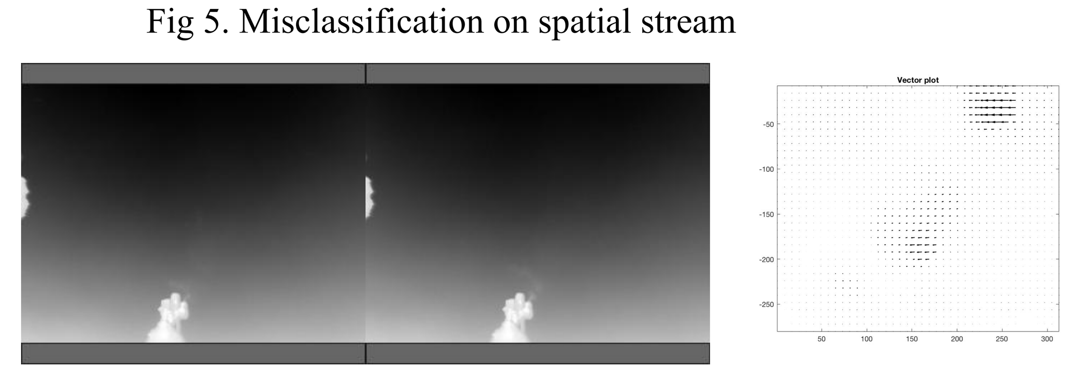


Fig 6. Misclassification on temporal stream

The left two images are from Class 1. In Class 1, the leak is really small. Then the optical flow method is sensitive to motion noise. Thus, from the calculation, we find that not only the area covering the plume has velocity filed. Then the prediction will be larger.

Discussion & Conclusion

- ✓ Generally, spatial stream can achieve high accuracy, because the spatial information can capture important features, including texture, plume area, pattern. Temporal stream can supplement more motion information. Thus the integration of the two stream model can have much higher accuracy.
- ✓ Since the plume is not a rigid body, optical flow cannot calculate the velocity field very accurately. Moreover, when the leak size is low, the optical flow is sensitive to the motion noise in the images. Thus only temporal stream cannot compete with spatial stream.
- ✓ The reasons why there are misclassifications are (1) affect of wind on the coverage and the shape of the plume, (2) accuracy of the optical flow calculation (3) sometimes the plume is covered by the cloud (4) all the leaks in the dataset are very small in the real world.
- ✓ The leaks from Class 1 were usually misclassified. Other classes all have high true positive rates.

Future Work

- ✓ Improve optical flow method. Find an optical flow method which is best for non-rigid body.
- ✓ Try different optical flow stacking methods. We can stack the two flow channels of L consecutive frames to form a total of 2L input channels, where L can be 1, 3, 5, and 10.
- ✓ Collect more data and test the algorithm on different videos.
- ✓ Instead of treating the problem as classification problem, due to the small range of the leak size in the dataset, I may solve the problem as a regression problem in order to treat the mislabel problem differently and increase the accuracy.
- ✓ Try different fusion method, like weighted sum of the loss.
- ✓ Integrate the wind speed and orientation into the model.
- ✓ Use pre-trained model in spatial stream. The spatial model is indeed image classification architecture. Thus we can build the temporal model on the existing advanced large-scale recognition methods, like VGG, ResNet and Inception.
- ✓ Try Long-term Recurrent Convolutional Networks to estimate the leak size.

References

[1] Brandt, A. R., Heath, et al. (2014). Methane leaks from North American natural gas systems. Science, 343(6172), 733-735.
[2] Miller, S. M., Wofsy, S. C., et al. (2013). Anthropogenic emissions of methane in the United States. Proceedings of the National Academy of Sciences,110(50), 20018-20022.
[3] Zavala-Araiza, D., Lyon, D. R., Alvarez, R. A., Davis, K. J., Harriss, R., Herndon, S. C., ... & Marchese, A. J. (2015). Reconciling divergent estimates of oil and gas methane emissions. Proceedings of the National Academy of Sciences, 112(51), 15597-15602.
[4] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 1725-1732).
[5] Simonyan, K., & Zisserman, A. (2014). Two-stream convolutional networks for action recognition in videos. In Advances in neural information processing systems (pp. 568-576).
[6] Sun, D., Roth, S., & Black, M. J. (2010, June). Secrets of optical flow estimation and their principles. In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on (pp. 2432-2439). IEEE.
[7] Farnebäck, G. (2003). Two-frame motion estimation based on polynomial expansion. Image analysis, 363-370.

Acknowledgements

Thank you all the instructors and TAs throughout the course, especially Rishi. You are all devoted, patient and knowledgeable. We would like to thank Professor Adam Brandt and Arvind P. Ravikumar for their help and support throughout the project.