# Fire Detection based on Convolutional Neural Network

Sungjae Park

# Outline

1. Introduction
2. Background of CNN
3. Methods
4. Results
5. Conclusion

# Introduction

According to NFPA report 2017,

Total 1,319,500 fires only in the US, resulted 34,000 fire deaths, 14,670 fire injuries, and $23 billion property loss

Fire injury and death occurred every 36 minutes and 2 hours and 34 minutes respectively.

# Fire detection using a surveillance camera

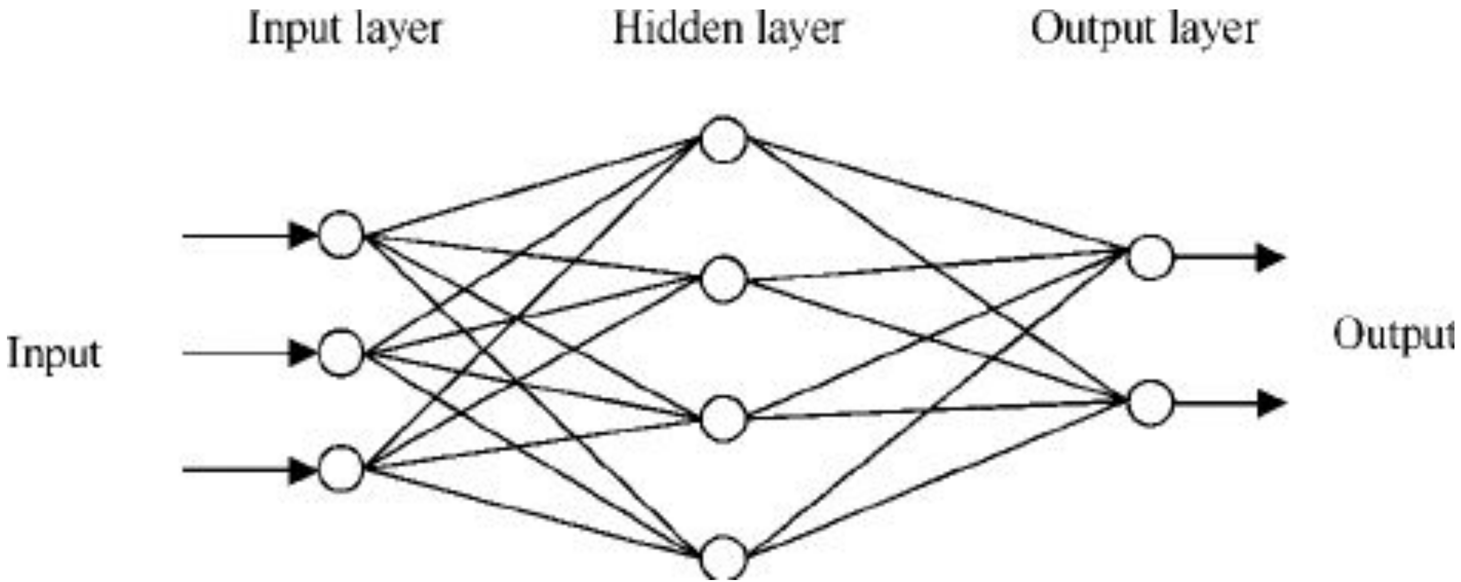
Fire detection using sensors including temperature detection and smoke detection

-> Disadvantage: lack of information on

1) location of fire outbreak
2) the direction of smoke distribution
3) intensity of the fire

Surveillance devices can remedy the disadvantage of sensors detection method.

This explains about weight, Layer, node(neuron) and backpropagation in CNN.



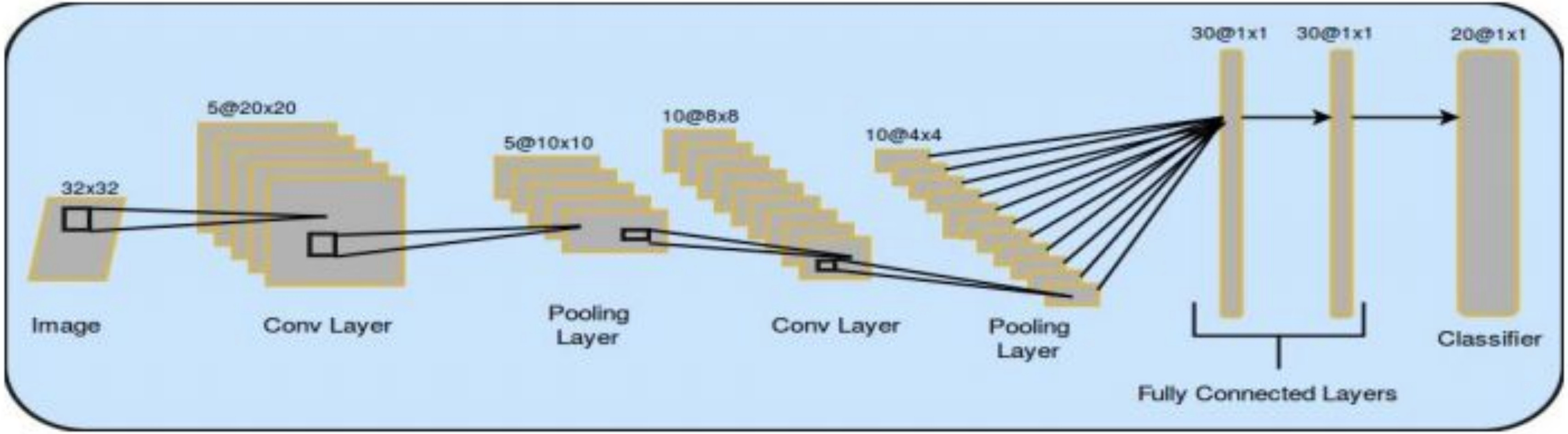
# CNN (Convolutional Neural Network)

CNN is one of artificial neural network (ANN) and is widely used in the image classification.

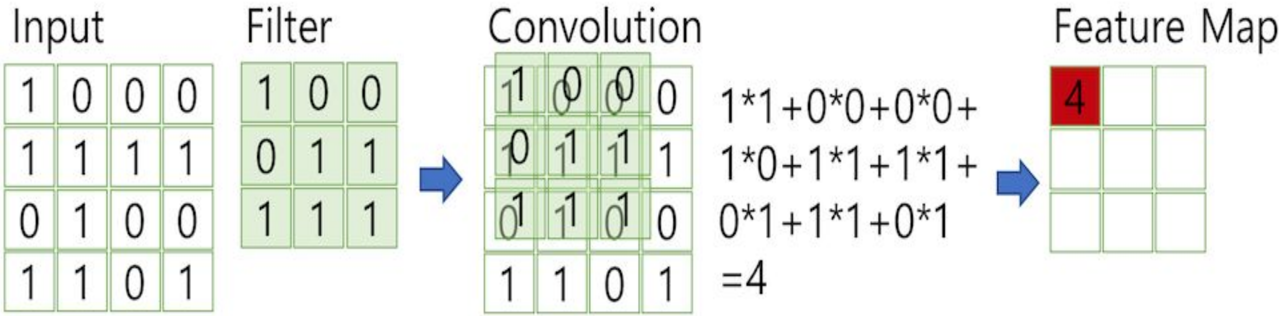
Better method than classical ANN in larger and complex models by having a spatially independent features.

CNN has multiple layers:

- Convolution layer
- Stride and padding
- Pooling layer
- Fully Connected Layer

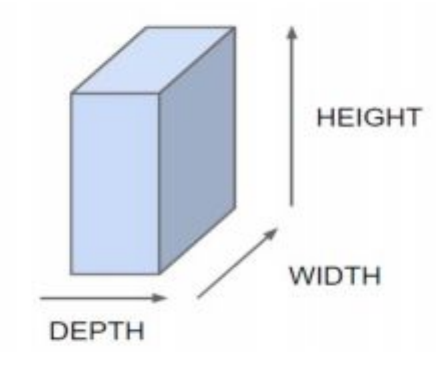


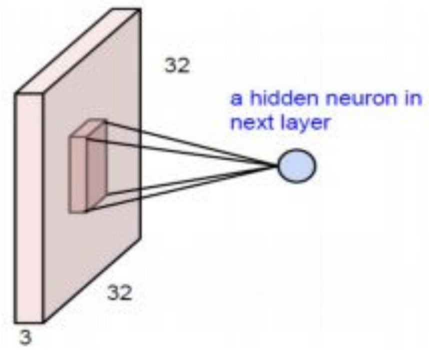

- Filter
- Convolution

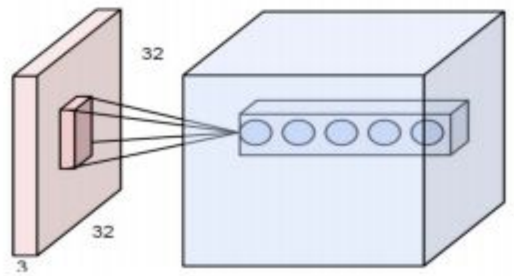


# Convolution Layer

- First layer in CNN to extract features from an input image.
- 3D input: width & height - pixels, depth - 3 RGB channel (Red, Green, Blue)
- Local connectivity concept: connect each neuron to only a local region of the input volume.
    - Reduce the number of parameters
    - Less weight connections
- Better method: fixed weights of local connection + apply the whole neurons of the next layer
    - Useful: adding more layers after input layer with different filters.



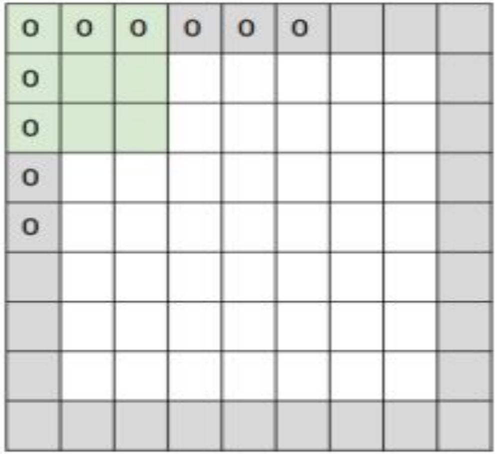



# Stride & zero-padding

- Control the volume of output by moving the filter.
- Stride 1: filter moves 1 pixel at a time ➔ Output: 5x5 volume.
- Stride 2: filter moves 2 pixels at a time ➔ Output: 3x3 volume.
- Issues with stride 3: receptive fields do not fit on the input volume. (Lose information!)
- Zero-padding: add zeros around the input borders
  - Keep the original input volume
  - Keep as many information as we can

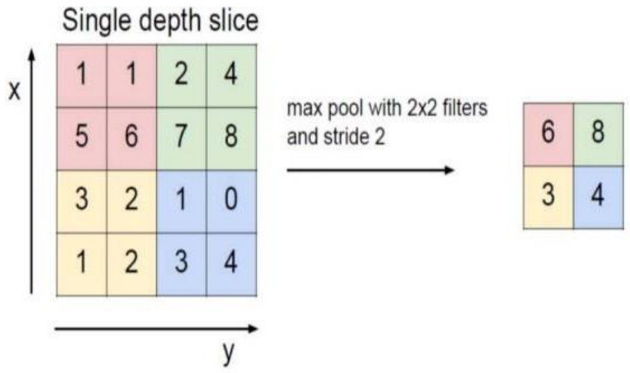
Stride 1




Zero-padding

# Pooling Layer

- Reduce the number of parameters and complexity for further layers
- Downsampling: Reduce the size of dimensions in the map and maintain the important information.
- Max pooling: the most common method
    - Max pooling with 2x2 filters and stride 2 - common size
        - Takes the largest values in each sub regions
        - Reduce the size to 2x2 volume + important information



# Fully Connected Layer

- Compared to neural network;
  each node is connected to every node

- Outputs of layers are high-level features of
  the input image

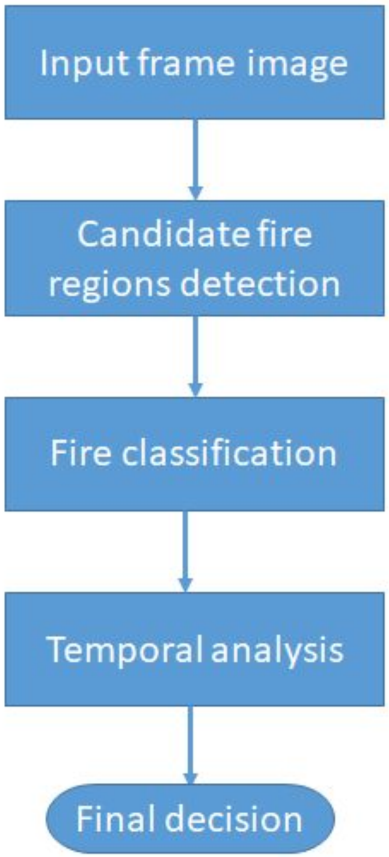
- Use high-level features for classifying the input image
  into various classes

- Takes a long time in training due to numerous connections

- Eliminate the number of nodes for reducing training times

# Methods

Dung and Ro (2018) : a fire classifier with a cascade model for fire detection algorithm using surveillance camera



# Methods - Detection of candidate fire

To detect candidate fire regions, they used RGB color map and the flickering energy map to estimate every image pixel because it's important factors to detect possible fire regions.

Color map

$$R \geq G \geq B \quad (1)$$

$$R > R_T \quad (2)$$

$$S > (255 - R)\, S_T / R_T \quad (3)$$

Flickering energy map

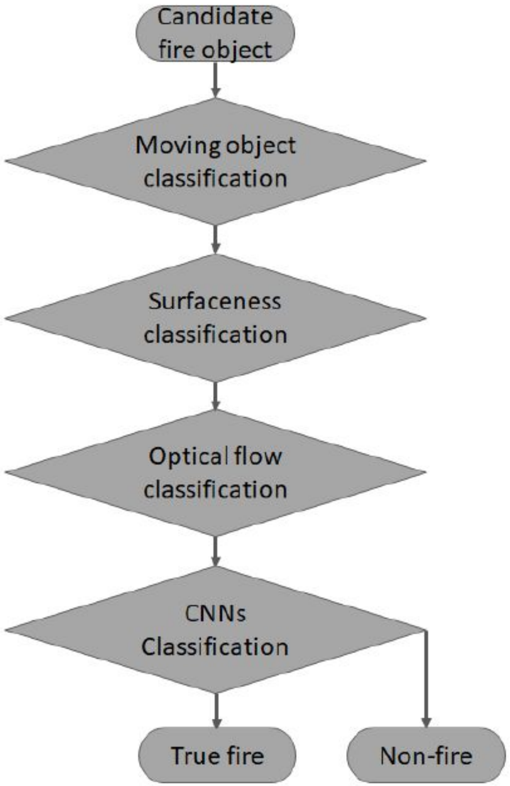
$$dI(t) = I(t) - I(t\text{-}1) \quad (4)$$

# Methods - Fire classification

Introduced by Dung and Ro 2018

1) Moving object classification
   - Calculate moving distance for the period of time

2) Surface feature classification
   - Reduce the turbulence; distinguish the surfaces of the fire and non-fire objects



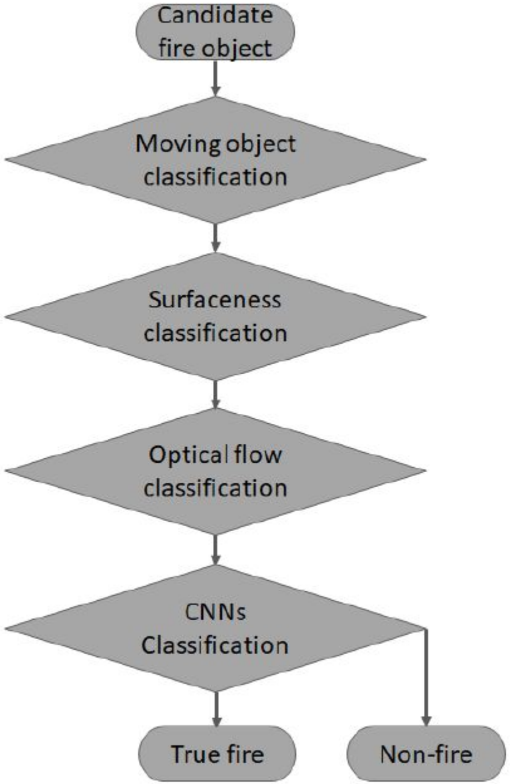
# Methods - Fire classification(Continued)

Introduced by Dung and Ro 2018

3) Optical flow classification
   - Since part of flames could possibly move in an unexpected direction

   - Every angles of optical flow vectors were calculated

   - Move on to the next layer, the CNN classifier



# Methods

Muhammad et al., (2018) : a cost-effective fire detection CNN architecture for surveillance videos.

Their model is inspired by GoogleNet mainly focusing computational complexity and accuracy compared to other networks such as AlexNet.

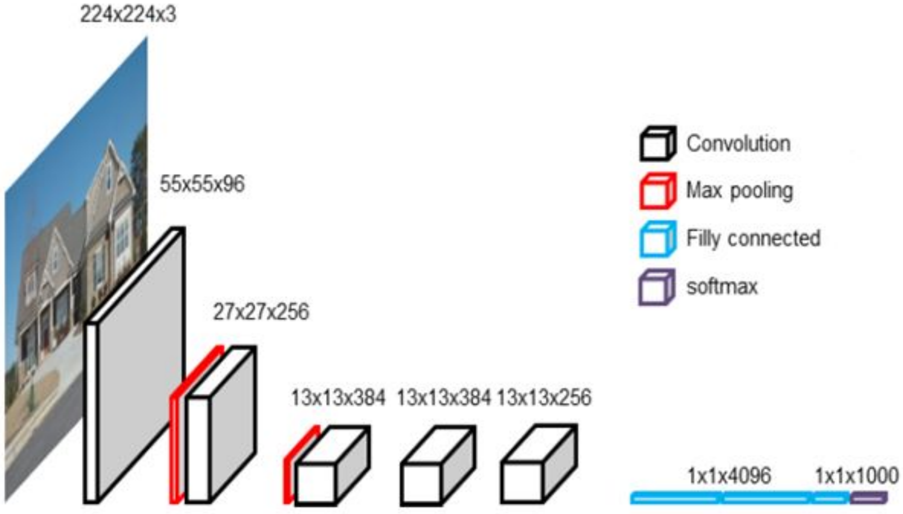
# Methods - AlexNet

Developed by Krizhevsky 2012

- One of the most influential architecture and Examined variations in CNN framework

- Contains 5 convolutional layers, 3 fully connected layers and soft max

- Soft Max is activation function.



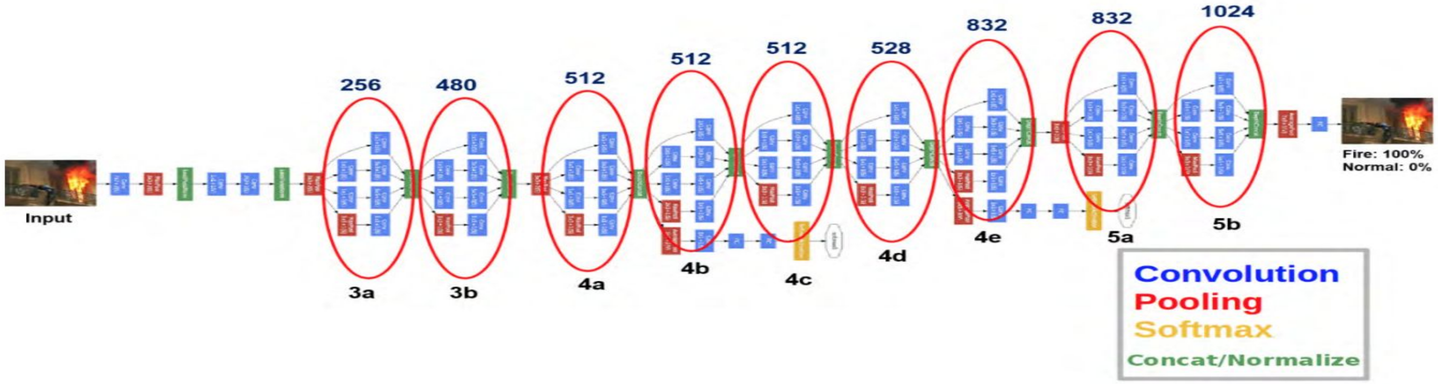
# Methods - GoogleNet

Developed by Google 2014

- Better classification accuracy, small sized model, and suitability of implementation on other hardware architectures having memory constraints

- Consists of 22 layers with 2 main convolutions, 4 max pooling, one average pooling, 9 inception modules

-



| type | patch size/ stride | output size | depth | #1×1 | #3×3 reduce | #3×3 | #5×5 reduce | #5×5 | pool proj | params | ops |
|---|---|---|---|---|---|---|---|---|---|---|---|
| convolution | 7×7/2 | 112×112×64 | 1 | | | | | | | 2.7K | 34M |
| max pool | 3×3/2 | 56×56×64 | 0 | | | | | | | | |
| convolution | 3×3/1 | 56×56×192 | 2 | | 64 | 192 | | | | 112K | 360M |
| max pool | 3×3/2 | 28×28×192 | 0 | | | | | | | | |
| inception (3a) | | 28×28×256 | 2 | 64 | 96 | 128 | 16 | 32 | 32 | 159K | 128M |
| inception (3b) | | 28×28×480 | 2 | 128 | 128 | 192 | 32 | 96 | 64 | 380K | 304M |
| max pool | 3×3/2 | 14×14×480 | 0 | | | | | | | | |
| inception (4a) | | 14×14×512 | 2 | 192 | 96 | 208 | 16 | 48 | 64 | 364K | 73M |
| inception (4b) | | 14×14×512 | 2 | 160 | 112 | 224 | 24 | 64 | 64 | 437K | 88M |
| inception (4c) | | 14×14×512 | 2 | 128 | 128 | 256 | 24 | 64 | 64 | 463K | 100M |
| inception (4d) | | 14×14×528 | 2 | 112 | 144 | 288 | 32 | 64 | 64 | 580K | 119M |
| inception (4e) | | 14×14×832 | 2 | 256 | 160 | 320 | 32 | 128 | 128 | 840K | 170M |
| max pool | 3×3/2 | 7×7×832 | 0 | | | | | | | | |
| inception (5a) | | 7×7×832 | 2 | 256 | 160 | 320 | 32 | 128 | 128 | 1072K | 54M |
| inception (5b) | | 7×7×1024 | 2 | 384 | 192 | 384 | 48 | 128 | 128 | 1388K | 71M |

# Training CNNs

AlexNet & GoogleNet

- Total number of images used in training experiments is 63,690
- Using 20% data of the whole dataset for training and remaining for testing
- Testing dataset 1: 31 videos; 14 fire videos and 17 normal videos without fire
- Testing dataset 2: 226 images; 119 images of fire and 107 images of non-fire

# Result

1. CaffeNet

   - Easy to detect the fire with high precision of CNNs image classifier

   - The resulting accuracy of the trained CNN classifier was 98%

   - Time consumption:
     Less than 1 milliseconds using surface feature classification

   - Significant improvement over similar methods and fully suited for use in a real-time surveillance system

| Video No. | Description | Fire Detected/False Positive | | |
|---|---|---|---|---|
| | | Gunawaardena | Jessica Ebert | Our Algorithms |
| 1 | Outdoor day fire | Fire detected | Fire detected | Fire detected |
| 2 | Outdoor day fire | Fire detected | Fire detected | Fire detected |
| 3 | Outdoor day fire | Fire detected | Fire detected | Fire detected |
| 4 | Outdoor day fire | Fire detected | Fire detected | Fire detected |
| 5 | Human moving around fire | Fire detected | Fire detected | Fire detected |
| 6 | Outdoor day fire | Fire detected | Fire detected | Fire detected |
| 7 | Human moving around fire | Fire detected | Fire detected | Fire detected |
| 8 | Indoor night fire | Fire detected | Fire detected | Fire detected |
| 9 | Indoor night fire | Fire detected | Fire detected | Fire detected |
| 10 | Indoor night fire | Fire detected | Fire detected | Fire detected |
| 11 | Red moving human | No fire | No fire | No fire |
| 12 | Red moving human | No fire | No fire | No fire |
| 13 | Red moving human | False positive | False positive | No fire |
| 14 | Yellow moving human | No fire | No fire | No fire |
| 15 | Red moving human | False positive | False positive | No fire |

# Result - Continued

2. AlexNet & GoogleNet (Dataset 1)

   1) AlexNet

      - The resulting accuracy: 90.06%
      - The resulting accuracy after added variety to the weights: 94.39%
      - The false alarms rate was 9.22% and false negative score is 10.65%; problematic of fire detection

   2) GoogleNet

      - The resulting accuracy: 88.41%
      - The resulting accuracy after added variety to the weights: 94.43%
      - The false alarms rate are diminished from 0.11% to 0.054% and false negative score from 5.5% to 1.5%

| Technique | False Positives (%) | False Negatives (%) | Accuracy (%) |
|---|---|---|---|
| GoogleNet after fine tuning (FT) | 0.05 | 1.50 | 94.43 |
| GoogleNet before FT | 0.11 | 5.50 | 88.41 |
| AlexNet (after FT) | 9.07 | 2.13 | 94.39 |
| AlexNet (before FT) | 9.22 | 10.65 | 90.06 |

# Result - Continued

## 2. AlexNet & GoogleNet (Dataset 2)

1) AlexNet

- Before the fine-tuning:
  1) Precision: 0.85
  2) Recall: 0.92
  3) F-Measure: 0.88
- After the fine-tuning:
  1) Precision: 0.82
  2) Recall: 0.98
  3) F-Measure: 0.89

2) GoogleNet

- Before the fine-tuning:
  1) Precision: 0.86
  2) Recall: 0.89
  3) F-Measure: 0.88
- After the fine-tuning:
  1) Precision: 0.80
  2) Recall: 0.93
  3) F-Measure: 0.86

Recall - probability of selecting a relevant frame as key

Precision - relevancy of chosen key frames

F-measure - average of both recall and precision.

Recall and precision are complementary to each other

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$\mathrm{F-Measure} = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

# Conclusion

| Dung and Ro (2018) | Muhammad et al., (2018) |
|---|---|
| • Cascade Model<br>• RGB color map<br>• Moving object classification<br>• Surface feature classification<br>• Optical flow classification | • AlexNet<br>• GoogleNet |

# Questions ?

I sincerely thank my processor for senior seminar, Elena Machkasova, and reviewers for guidance and feedback.I wish to express my sincere gratefulness to my Academic Advisor Peter Dolan for providing me good advice.