BATCH LOSS REGULARIZATION IN DEEP LEARNING METHOD FOR AERIAL SCENE CLASSIFICATION

Yuanjun Huang (huangyuanjun76@126.com)

Prof. Xianbin Cao

Beihang University, Beijing, China

Beijing Laboratory for General Aviation Technology

Outline

1. Motivation

2. Problem Analysis

3. Proposed Model

4. Results and Conclusion

Motivation

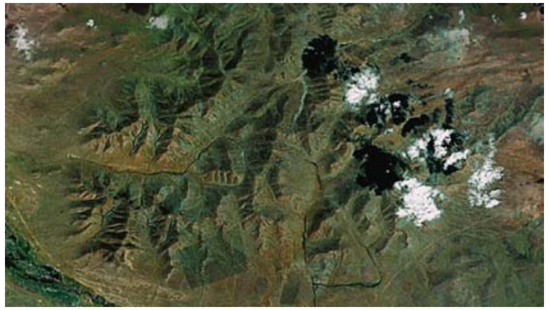
Aerial scene classification is a very essential element in Unmanned Aerial Vehicle (UAV) surveillance system. Also, it has wide applications in various computer vision tasks.

Unmanned Aerial Vehicle (UAV) based Aerial Scene Classification

Wide Applications

- Landform Analysis
- UAV Safety

- UAV Surveillance
- Data Classification



Unmanned Aerial Vehicle (UAV) based Aerial Scene Classification

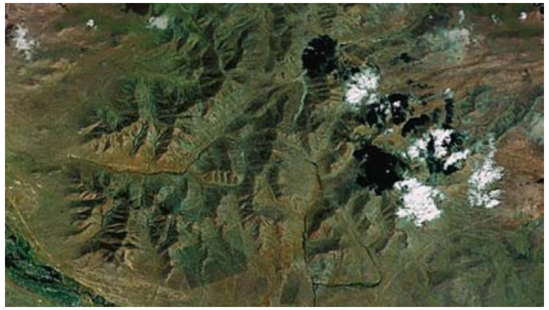
Huge quantities of aerial videos and images

Dynamic Scenes (Videos)

- Forest Fire
- Highway
- Street
- …

Aerial Scenes (Images)

- Parking
- Airport
- Mountain
- …



Unmanned Aerial Vehicle (UAV) based Aerial Scene Classification

Data samples

Dynamic Scenes

Iceberg collapse

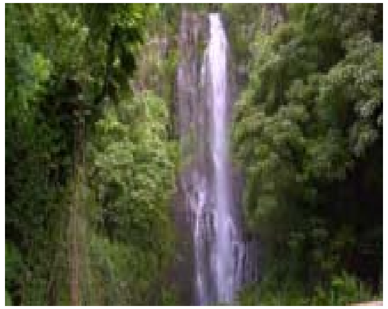
Volcano Eruption

Problem Analysis

Aerial scenes present large intra-class variations as well as small inter-class differences

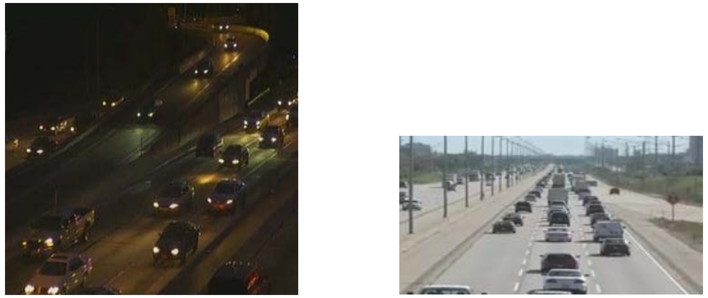
Unmanned Aerial Vehicle (UAV) based Aerial Scene Classification


Rush River


Waterfall


Fountain

- Similar objects in different classes


Highway

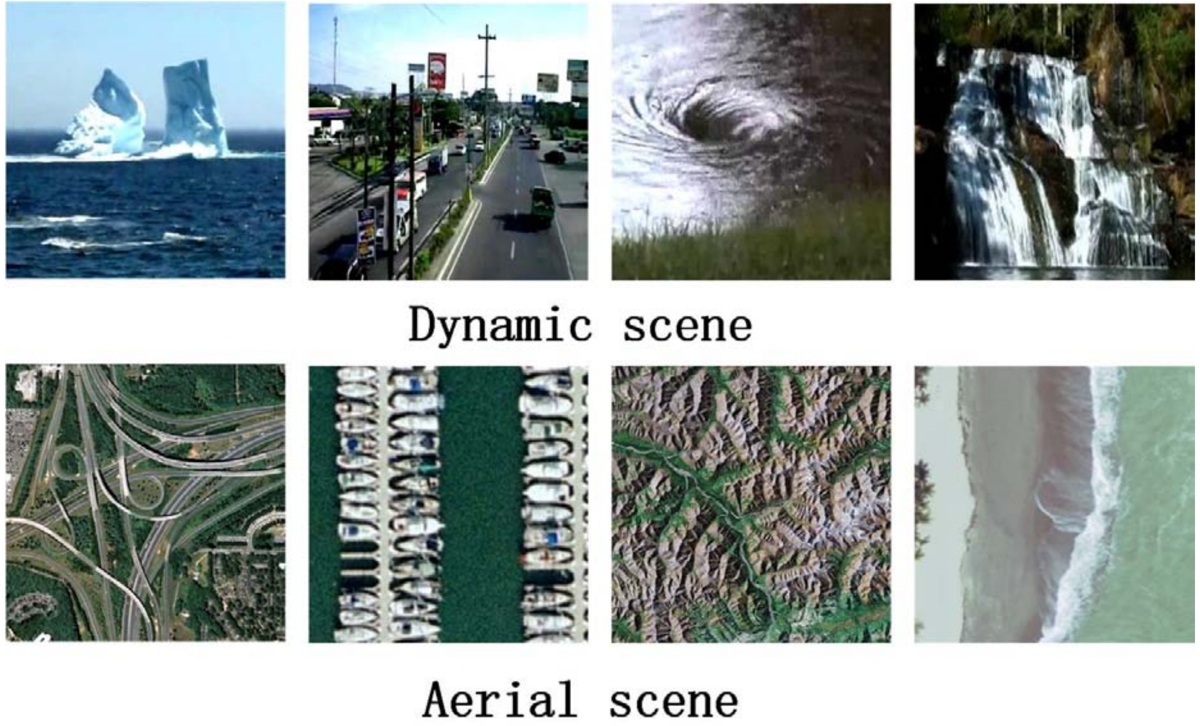
- Diversified appearance in single class

Small inter-class variations

Large intra-class variations

Difficulties in aerial scene classification

Unmanned Aerial Vehicle (UAV) based Aerial Scene Classification



Aerial scenes: Appearance clues for classification

Dynamic scenes: movement and appearance clues for classification

Unmanned Aerial Vehicle (UAV) based Aerial Scene Classification

Available methods:
Hand-crafted VS Deep-learned

Hand-crafted features:
- Appearance (SIFT,HOG,GIST,etc)
- Motion (Optical flow)
- Spatio-temporal (SOE,CSR)

Deep-learned features:
- Appearance (VGG, Googlenet, Alexnet, etc.)
- Spatio-temporal (C3D)

Summarize

The deep learning based methods have superior distinctive power over hand crafted features.

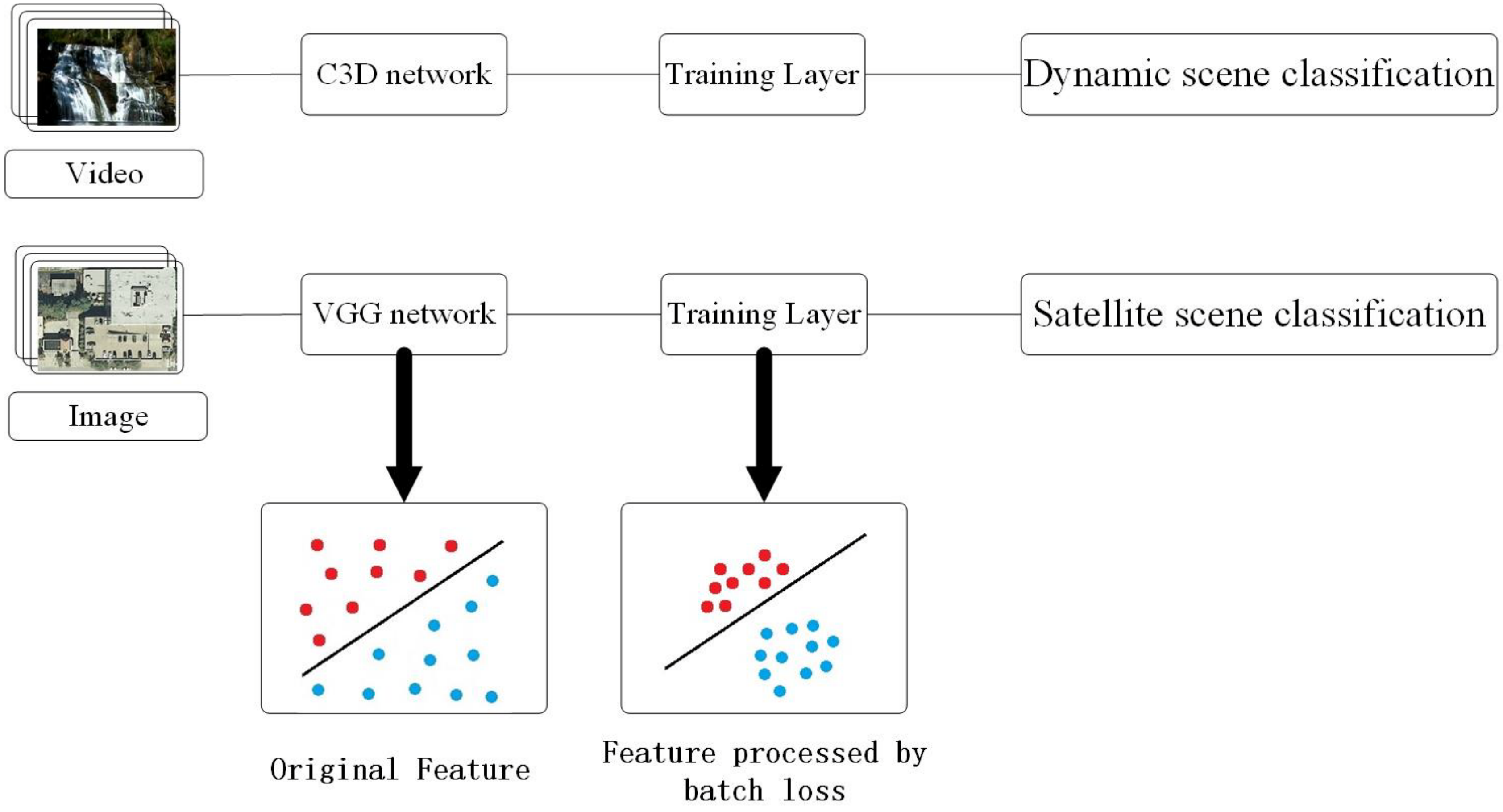
1. Deep learning based methods require large datasets for training.
2. Most deep learning methods concentrate on inter-class separable and ignore large intra-class variations.

Proposed Method

Focus on reducing large intra-class variations while keeping features separable

-----Batch loss regularization

Batch Loss Regularization in Deep Learning



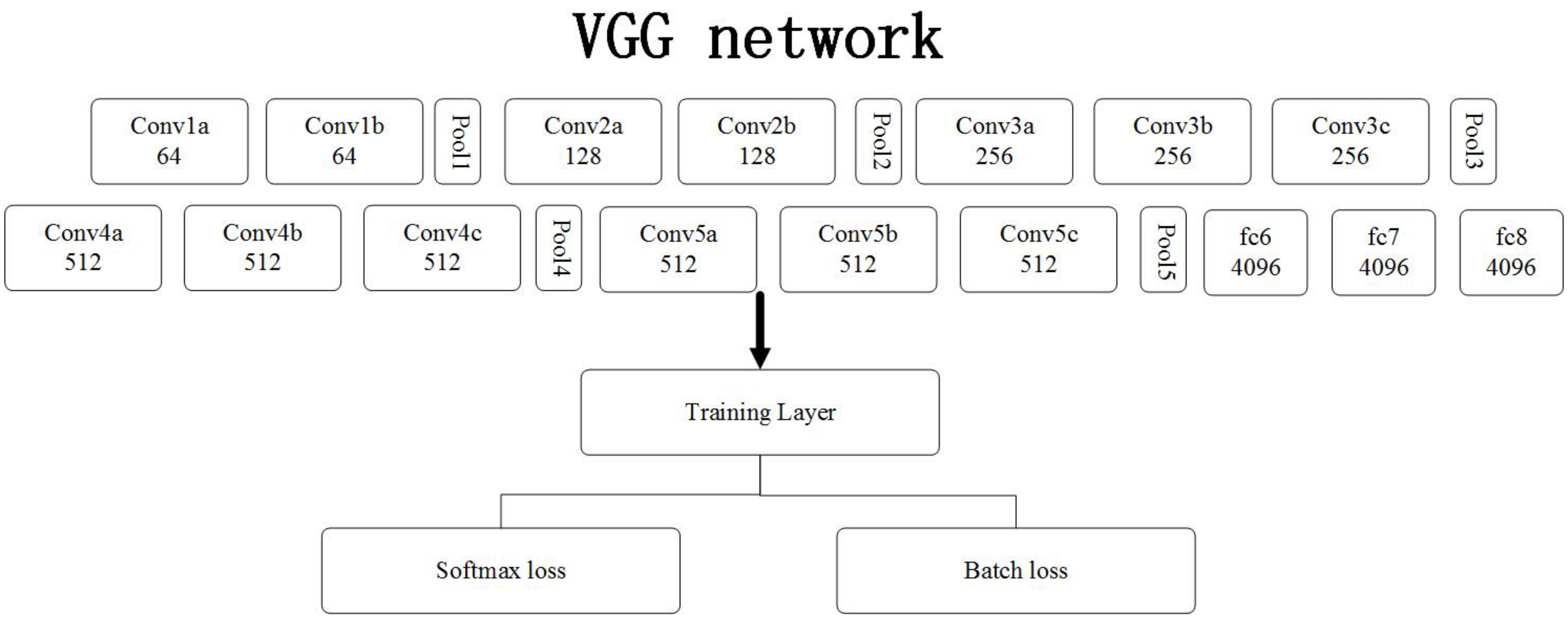
Batch Loss Regularization in Deep Learning

VGG & C3D network: Extract separable features

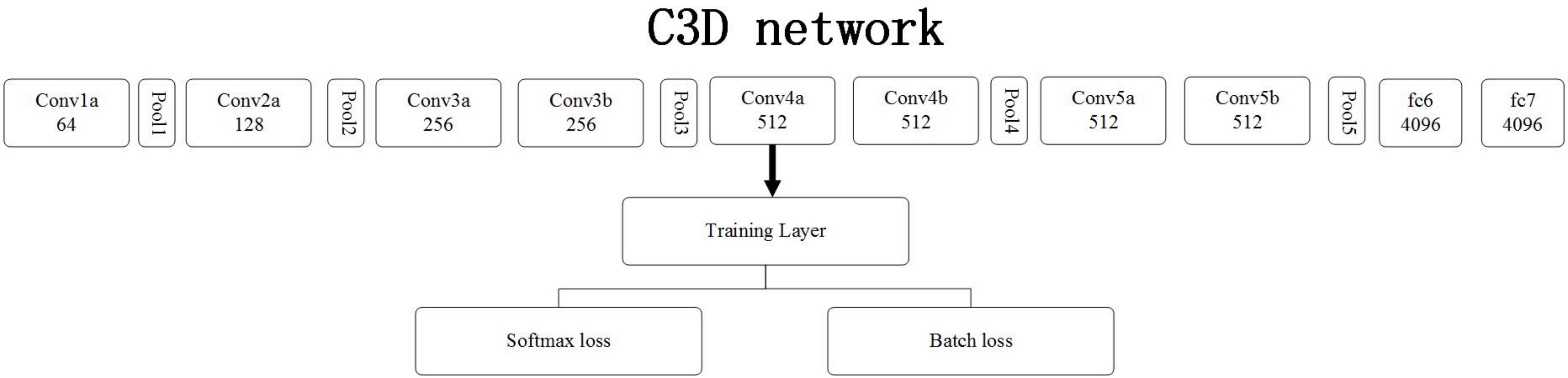
↓ ↑

Training layer: Reduce large intra-class variations while keeping features separable

Batch Loss Regularization in Deep Learning



Batch Loss Regularization in Deep Learning



Batch Loss Regularization in Deep Learning

Training Layer

$$\text{SoftmaxLoss} = -\sum_{i=1}^{M} \log \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}}$$

Keep features separable

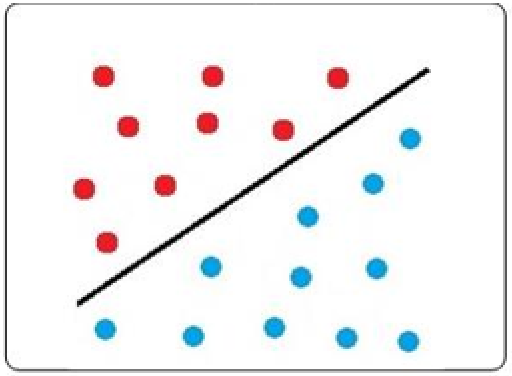
$$Batchloss = \frac{1}{2}\sum_{i=1}^{M}(x_i - B_c)^2$$

Reduce large intra-class variations

Batch Loss Regularization in Deep Learning

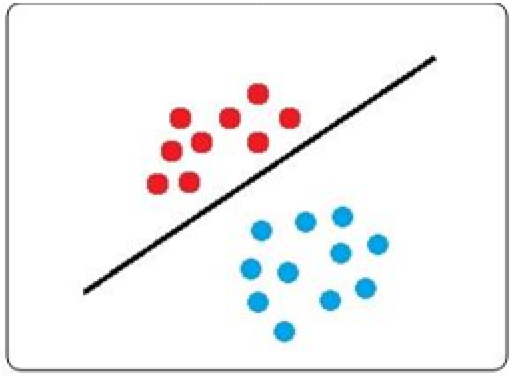
Training Layer

Softmax loss



Original Feature

Batch loss



Feature processed by batch loss

Results

Two applications:

Dynamic scene classification

Aerial image scene classification

Four datasets:

Dynamic scenes YUPENN dataset

Dynamic scenes Maryland dataset

Aerial scenes USGS dataset

Aerial scenes RS dataset

Dynamic scenes Datasets

The YUPENN dataset of dynamic scenes

The Maryland dataset of dynamic scenes

Batch Loss Regularization in Deep Learning

Dynamic scenes Datasets

Comparison Results on YUPENN Dataset

Method	Accuracy(%)
Chaos+GIST [1]	23
SFA[9]	85
SOE[11]	81
CSO[19]	86
BoSE[20]	96
CSR[26]	94
C3D[3]	97
C3D with batch loss	98

Comparison Results on Maryland Dataset

Method	Accuracy(%)
Chaos+GIST [1]	58
SFA[9]	60
SOE[11]	43
CSO[19]	68
BoSE[20]	78
CSR[26]	86
C3D[3]	78
C3D with batch loss	81

Batch Loss Regularization in Deep Learning

Aerial image scenes Datasets

The USGS dataset of aerial scenes

The RS dataset of aerial scenes

Batch Loss Regularization in Deep Learning

Comparison Results on USGS Dataset

Method	Accuracy(%)
BOVW[2]	76.8
SPM[2]	75.3
BOVW + co-occurrence kernel[2]	77.7
Color Gabor [2]	80.5
Color histogram [2]	81.2
Unsupervised learning [13]	81.7
Saliency-guided learning [27]	82.7
Wavelet BOVW [28]	87.4
Structural texture similarity [29]	86
Circle-structured BOVW [30]	86.6
Multifeature concatenation [31]	89.5
Pyramid-of-spatial-relatons [32]	89.1
MS-CLBP [33]	90.6
VGG [10]	91
VGG with batch loss	93

Batch Loss Regularization in Deep Learning

Comparison Results on RS Dataset

Method	Accuracy(%)
Bag of colors [31]	70.6
Tree of c-shapes [31]	80.4
Bag of SIFT [31]	85.5
Multifeature concatenation [31]	90.8
LTP-HF [34]	77.6
SIFT + LTP-HF + color [34]	93.6
MS-CLBP [33]	93.4
VGG [10]	95
VGG with batch loss	96

Future Work

Dynamic scene classification is more challenging than aerial image scene classification due to the lack of available training datas.

Thus, future work may include training the neural networks with limited data resources while keeping high accuracy performance.

Thank you

Q & A