FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
SCHOOL OF ENGINEERING

Fast Sample Generation with Variational Bayesian for Limited Data Hyperspectral Image Classification

July 26, 2018

AmirAbbas Davari, Hasan Can Özkan, Andreas Maier, Christian Riess

Pattern Recognition Lab

Friedrich Alexander University, Erlangen-Nürnberg

FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
SCHOOL OF ENGINEERING

Outline

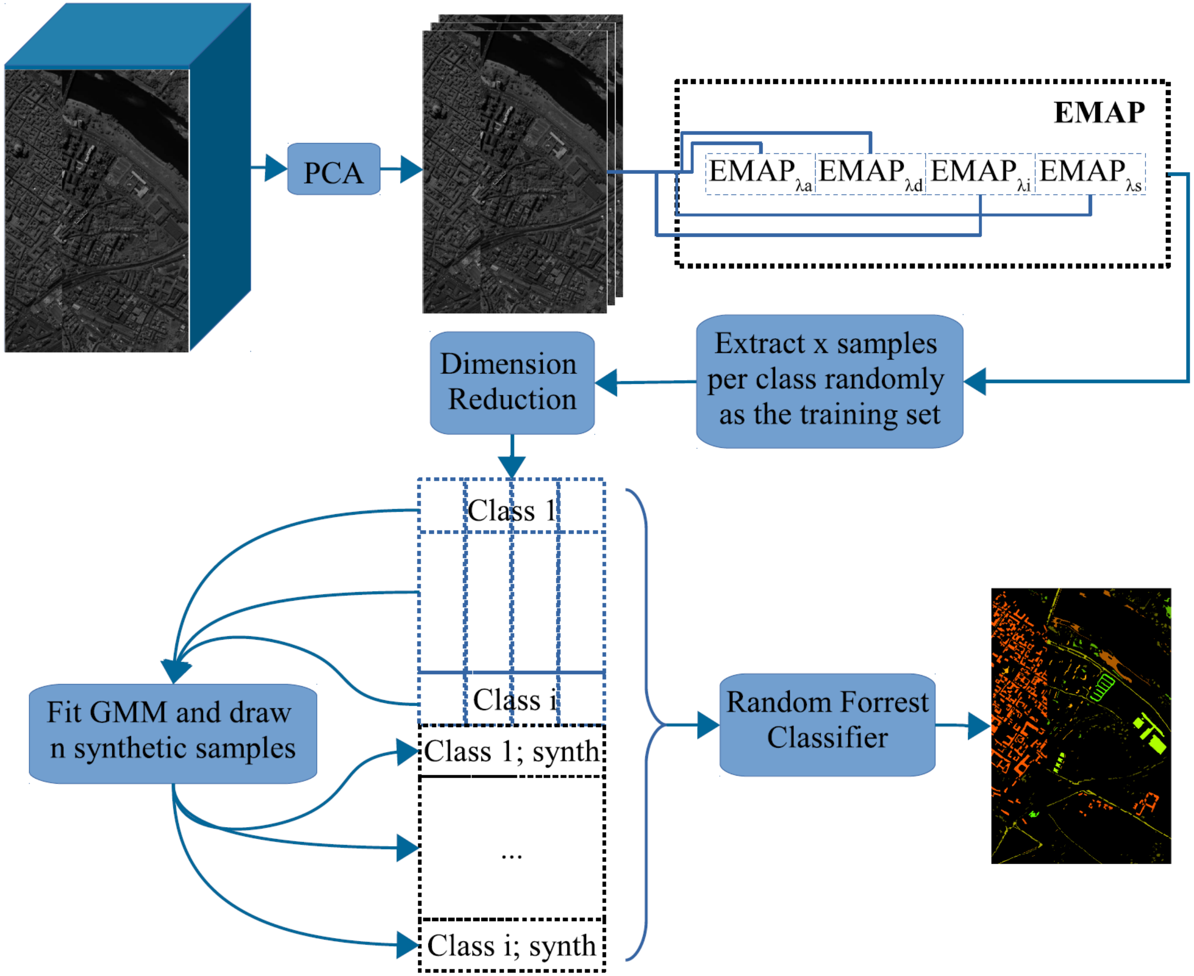
Introduction
- GMM-based synthetic HS data augmentation
- Variational Bayesian (VB)

Pipeline

Experimental Setup
- Dataset
- Classification
- Experimental Results

Conclusion

GMM-based Synthetic HS Data Augmentation [1]

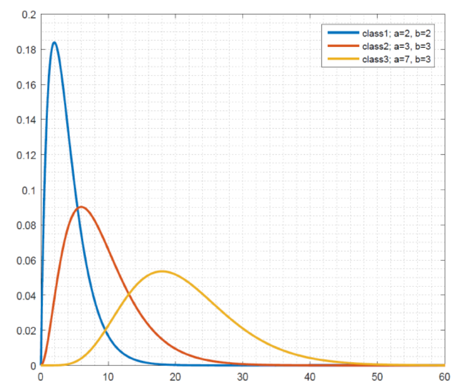


[1] Davari, Amirabbas, Erchan Aptoula, Berrin Yanikoglu, Andreas Maier, and Christian Riess. "GMM-Based Synthetic Samples for Classification of Hyperspectral Images With Limited Training Data." *IEEE Geoscience and Remote Sensing Letters 2018*

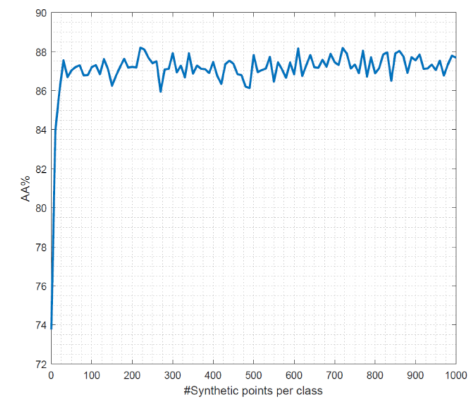
Datasets

Synthetic Dataset

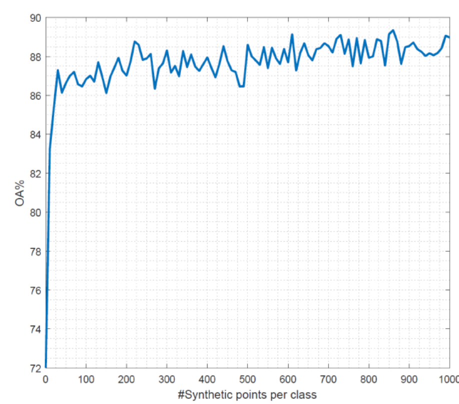
- Gamma Distribution
- Three different parameters represents Three classes
- 1000, 2000 and 3000 samples for blue, red and yellow classes
- 13 pixels per class were randomly drawn as training set



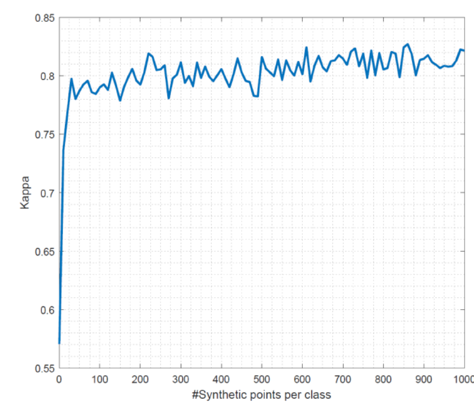
(a) Gamma distributions



(b) Average accuracy



(c) Overall accuracy



(d) Kappa

Variational Bayesian (VB)

- We would like to do maximum likelihood:

$$\theta_{MLE} = \underset{\theta}{\operatorname{argmax}} p(\mathcal{D}|\theta)$$

- $p(x|\theta)$ is the underlying distribution that we want to maximize.
- Why are we interested in $p(x|\theta)$ and not $p(x)$?
  - $p(x)$ is not expressive enough.

$$\begin{aligned} p(x|\theta) &= \int p(x, z|\theta) dz \\ &= \int p(x|\theta) p(z|x, \theta) dz \\ &= \int p(z|\theta) p(x|z, \theta) dz \end{aligned}$$

Variational Bayesian (VB)

- In a Gaussian Mixture Model:

$$p(x|\theta) = \int p(x, z|\theta)dz = \sum_{k=0}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)$$

$$\theta = \{\pi_k, \theta_k, \Sigma_k\}$$

- For the log likelihood, however:

$$\underset{\theta}{\operatorname{argmax}} \ln p(\mathcal{D}|\theta) = \sum_{n=1}^{N} \ln \sum_{k=0}^{K} \pi_k \mathcal{N}(x_n|\mu_k, \Sigma_k)$$

there is a logarithm inside the sum and there is no close form solution for that.

Variational Bayesian (VB)

$$\ln p(\mathbf{x}|\theta) = \int q(\mathbf{z}) \ln \left(p(\mathbf{x}|\theta) \frac{q(\mathbf{z})}{q(\mathbf{z})} \right) dz$$

$$= \underbrace{\int q(\mathbf{z}) \ln \frac{p(\mathbf{x}, \mathbf{z}|\theta)}{q(\mathbf{z})} dz}_{\mathcal{L}_{ELBO}(q,\theta)} + \underbrace{\int q(\mathbf{z}) \ln \frac{q(\mathbf{z})}{p(\mathbf{z}|\mathbf{x}, \theta)} dz}_{\mathbf{KL}(q(\mathbf{z})||p(\mathbf{z}|\mathbf{x},\theta))}$$

- Since $\mathbf{KL}(q(\mathbf{z})||p(\mathbf{z}|\mathbf{x},\theta)) \geq 0$; i.e. is non-negative, the first part is the lower bound of the log likelihood (evidence) i.e. the evidence lower bound or ELBO in short.

$$\mathcal{L}_{ELBO}(q,\theta) = \int q(\mathbf{z}) \ln \frac{p(\mathbf{x},\mathbf{z}|\theta)}{q(\mathbf{z})} dz \leq \ln p(\mathbf{x}|\theta)$$

$$=> \operatorname*{argmin}_q \mathbf{KL}(q(\mathbf{z})||p(\mathbf{z}|\mathbf{x},\theta)) = \operatorname*{argmax}_q \mathcal{L}_{ELBO}(q,\theta)$$

Variational Bayesian (VB)

- Assume the responsibility $r_{nk}(\theta)$

$$r_{nk}(\theta) \equiv p(z|\mathbf{x}_n, \theta)$$

- It turns out that $r_{nk}(\theta)$ is the optimal solution to

$$\underset{q}{\operatorname{argmin}} \, \mathbf{KL}(q(\mathbf{z})||p(\mathbf{z}|\mathbf{x}, \theta)) = \underset{q}{\operatorname{argmax}} \, \mathcal{L}_{ELBO}(q, \theta)$$

- Further, after some math it turns out that:

$$\underset{\theta}{\operatorname{argmax}} \, \mathcal{L}_{ELBO}(q, \theta) = \underset{\theta}{\operatorname{argmax}} \, \mathbf{E}_{q(\mathbf{z})} \left[\ln p(\mathbf{x}, \mathbf{z}|\theta)\right]$$

- The learning is an EM like procedure. Fix θ and maximize $r_{nk}(\theta)$, and then fix $r_{nk}(\theta)$ and maximize θ .

Datasets

Pavia Centre Scene Dataset:

- acquired by the ROSIS sensor
- 610 * 340 pixels
- geometrical resolution of 1.3 m
- 103 spectral bands
- We used the first four of its PCs which contained 99.16% of the total variance

#	Class	Samples
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947

Datasets

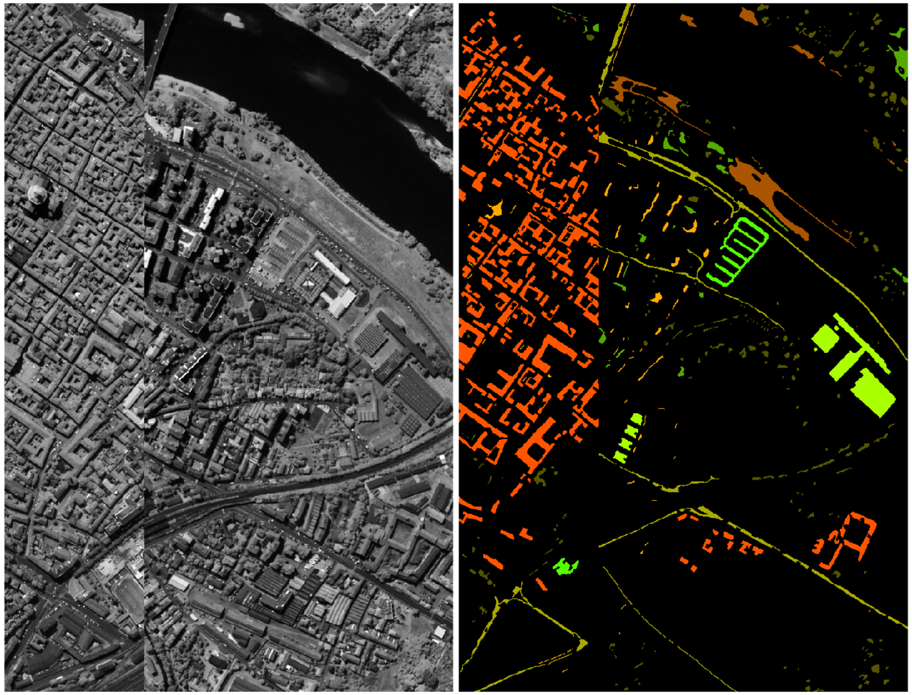
Salinas Valley Scene Dataset:

- acquired by the AVIRIS sensor
- 512 * 217 pixels
- geometrical resolution of 3.7 m
- 204 spectral bands
- We used the first four of its PCs which contained 99.68% of the total variance

#	Class	Samples
1	Brocli_green_weeds_1	2009
2	Brocli_green_weeds_2	3726
3	Fallow	1976
4	Fallow_rough_plow	1394
5	Fallow_smooth	2678
6	Strubble	3959
7	Celery	3579
8	Grapes_untrained	11271
9	Soil_vinyard_develop	6203
10	Corn_senesced_green_weeds	3278
11	Lettuce_romaine_4wk	1068
12	Lettuce_romaine_5wk	1927
13	Lettuce_romaine_6wk	916
14	Lettuce_romaine_7wk	1070
15	Vinyard_untrained	7268
16	Vinyard_vertical_trellis	1807

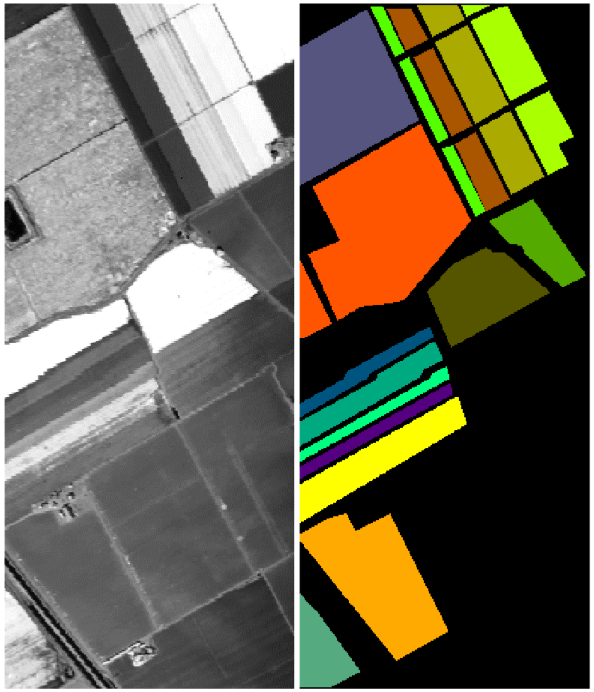
Datasets

Pavia Centre



First PC Classes

Salinas Valley

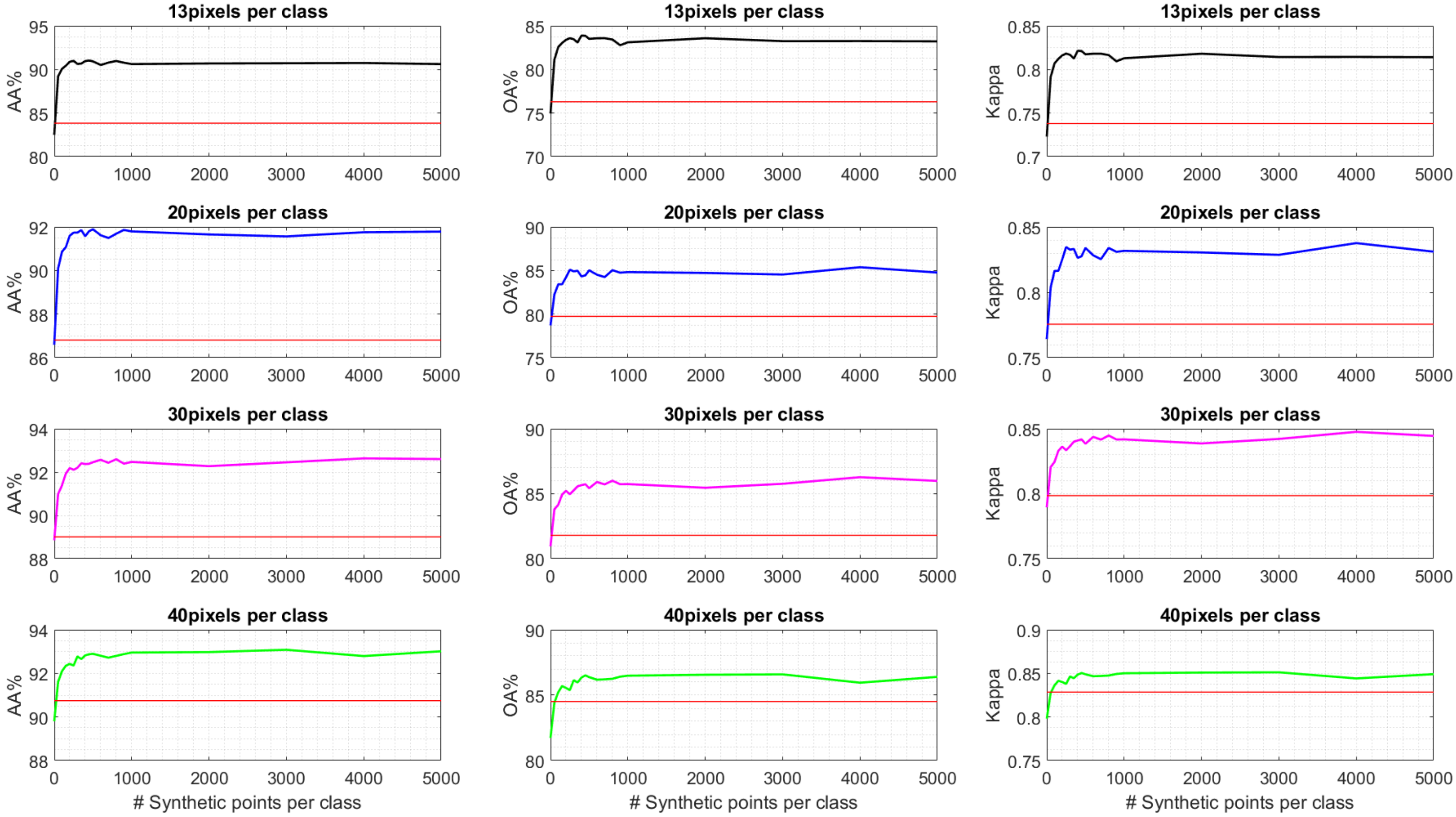


First PC Classes

Classification

- Classifier:
  - random forest with 100 trees as height and square root of number of features as depth.
- Training set: 13, 20, 30 and 40 pixels per class were randomly selected from the image as separate training sets.
- For each experiment, the random selection of the training set was repeated 25 times and the average of the overall accuracy, average accuracy and kappa statistics were calculated and reported.

Results: Salinas-PCA-EMAP-PCA



Results: Salinas-PCA-EMAP-NWFE

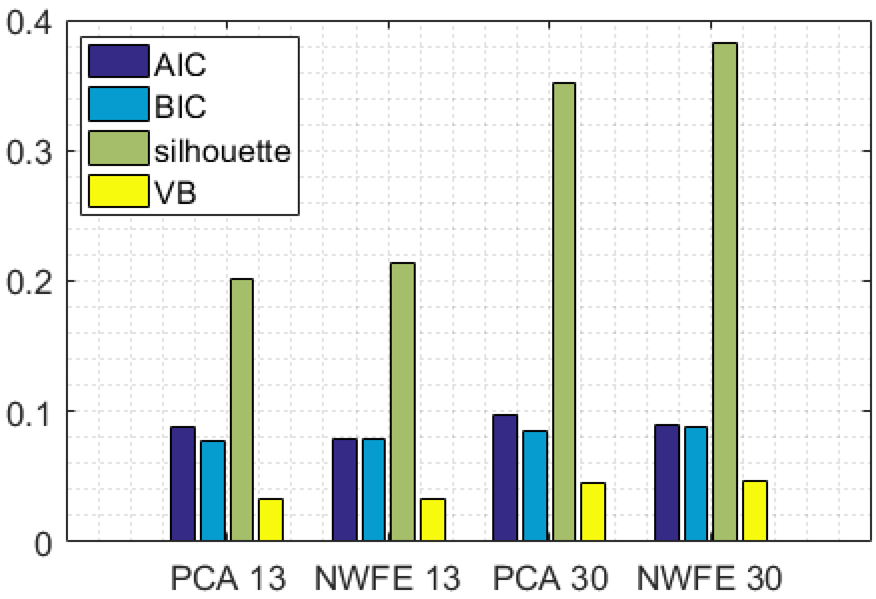
Algorithm	AA % (± SD)	OA % (± SD)	Kappa (± SD)
13 pix/class			
EMAP	83.84 (±2.06)	76.30 (±2.74)	0.7380 (±0.0292)
EMAP-NWFE	88.68 (±1.20)	80.42 (±2.34)	0.7838 (±0.0247)
EMAP-NWFE-Synth	**92.86** (±0.72)	**86.19** (±1.79)	**0.8468** (±0.0195)
20 pix/class			
EMAP	86.81 (±1.63)	79.74 (±2.56)	0.7756 (±0.0269)
EMAP-NWFE	90.56 (±1.26)	82.26 (±2.62)	0.8038 (±0.0280)
EMAP-NWFE-Synth	**93.70** (±0.46)	**87.30** (±1.27)	**0.8590** (±0.0139)
30 pix/class			
EMAP	89.01 (±1.10)	81.80 (±2.30)	0.7985 (±0.0248)
EMAP-NWFE	92.25 (±0.82)	84.76 (±2.38)	0.8314 (±0.0256)
EMAP-NWFE-Synth	**94.25** (±0.47)	**88.20** (±1.48)	**0.8690** (±0.0162)
40 pix/class			
EMAP	90.75 (±0.86)	84.52 (±1.76)	0.8285 (±0.0192)
EMAP-NWFE	93.29 (±0.41)	86.09 (±1.86)	0.8462 (±0.0200)
EMAP-NWFE-Synth	**94.60** (±0.41)	**88.92** (±1.08)	**0.8768** (±0.0119)

FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
SCHOOL OF ENGINEERING

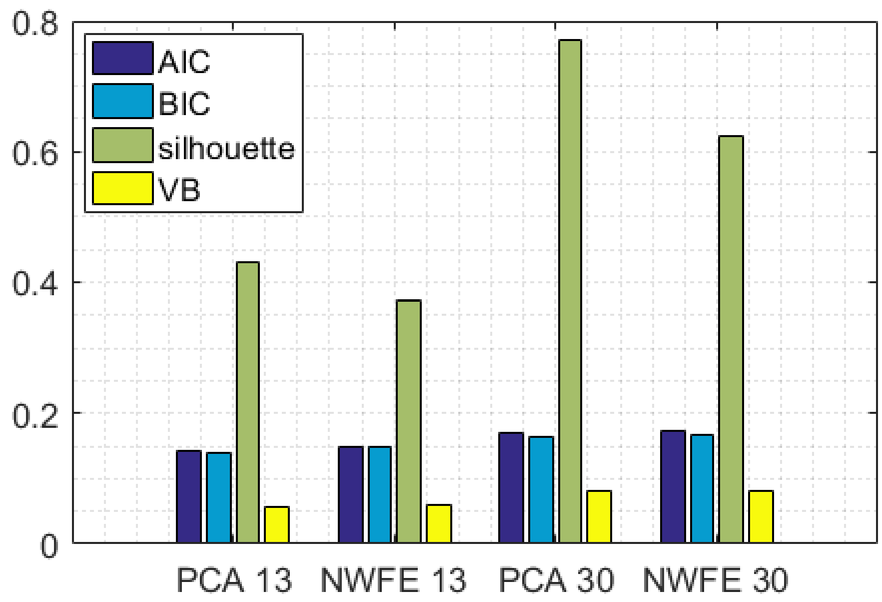
Results: Pavia Centre, VB vs. the other methods

Algorithm	pix per class	AA% (±SD)	OA% (±SD)	Kappa (±SD)	Run time (s) (±SD)
PCA					
AIC	13	85.17 (±1.21)	93.39 (±1.44)	0.9072 (±0.0197)	0.0877 (±0.0115)
	30	88.52 (±0.71)	94.68 (±0.51)	0.9252 (±0.0070)	0.0978 (±0.0054)
BIC	13	85.50 (±1.14)	93.67 (±0.82)	0.9110 (±0.0114)	0.0781 (±0.0017)
	30	88.23 (±1.06)	94.81 (±0.46)	0.9270 (±0.0064)	0.0856 (±0.0025)
avg. silhouette	13	85.10 (±1.27)	93.64 (±1.03)	0.9105 (±0.0142)	0.2013 (±0.0082)
	30	87.61 (±1.14)	94.68 (±0.32)	0.9251 (±0.0045)	0.3521 (±0.0223)
gap	13	83.14 (±1.87)	92.23 (±1.00)	0.8911 (±0.0138)	16.4766 (±0.1400)
	30	85.87 (±2.62)	93.54 (±1.03)	0.9091 (±0.0145)	35.4273 (±0.2511)
VB	13	85.60 (±0.62)	93.52 (±0.40)	0.9090 (±0.0055)	0.0324 (±0.0030)
	30	89.14 (±0.46)	95.15 (±0.42)	0.9317 (±0.0059)	0.0450 (±0.0029)

Results:

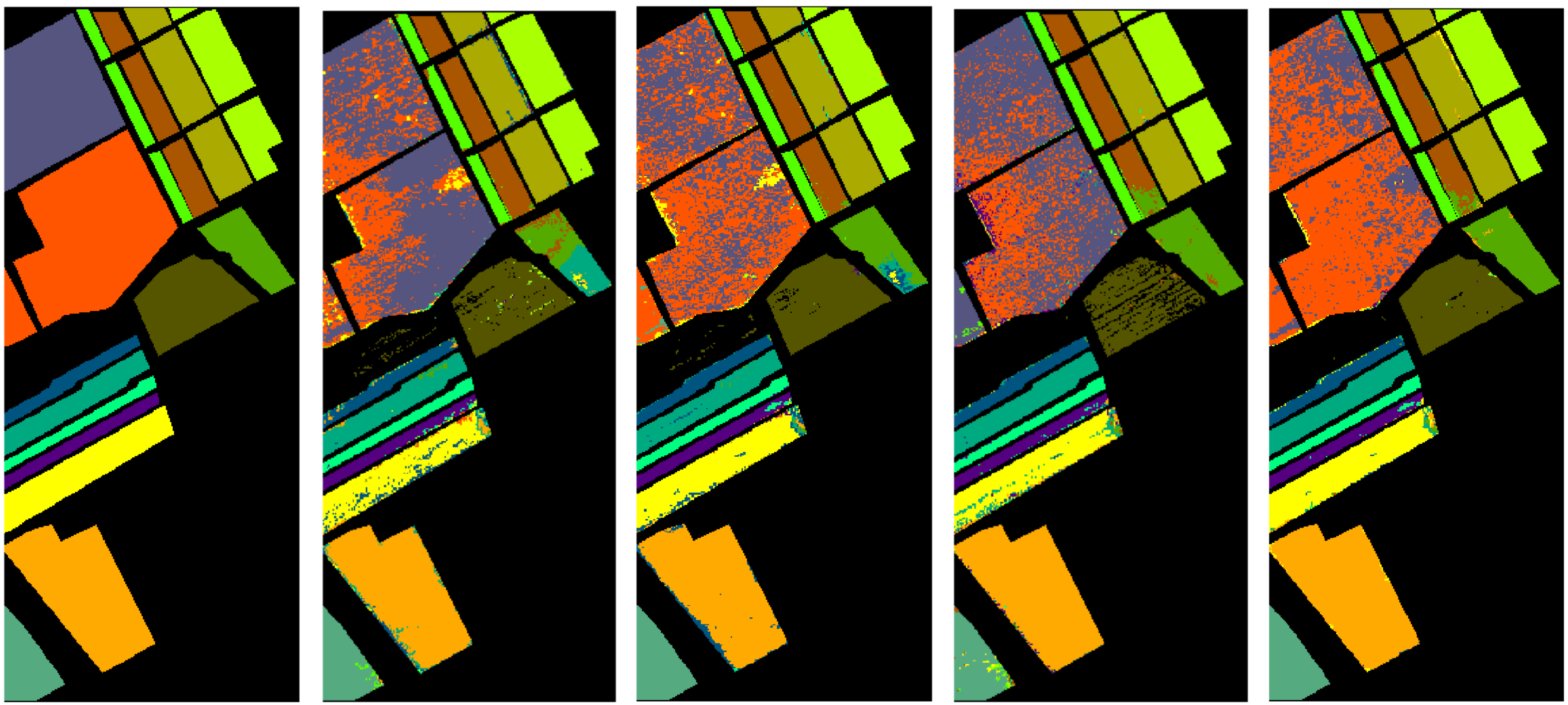


Pavia Centre



Salinas

Results: Salinas



Ground Truth | EMAP-PCA | EMAP-PCA-Synth | EMAP-NWFE | EMAP-NWFE-Synth

Conclusion

- VB yields similar, if not better, classification performance compared to the other methods.
- The results using VB are generally more consistent as well.
- More importantly, Variational Bayesian does not need the clustering algorithm to be executed in advance.
- VB is very memory efficient and drastically reduces the computational cost

FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
SCHOOL OF ENGINEERING

Thank you for your attention!