Comparing Supervised Learning Algorithms for Image Classification and Segmentation

Calarina Muslimani[1], Daria Garkavtseva[2], Ekaterina Rapinchuk[3]

University of Maryland, College Park[1], University of Massachusetts, Amherst[2], Michigan State University[3]

Background

- **Hyperspectral images** process visual information across the **electromagnetic spectrum**
  - Contains hundreds of channels
  - Various thematic applications such as ecological and hydrological sciences

Objectives

Hyperspectral Image Segmentation:

Classify **Pavia Centre and Pavia University** images, pixel-by-pixel, into classes

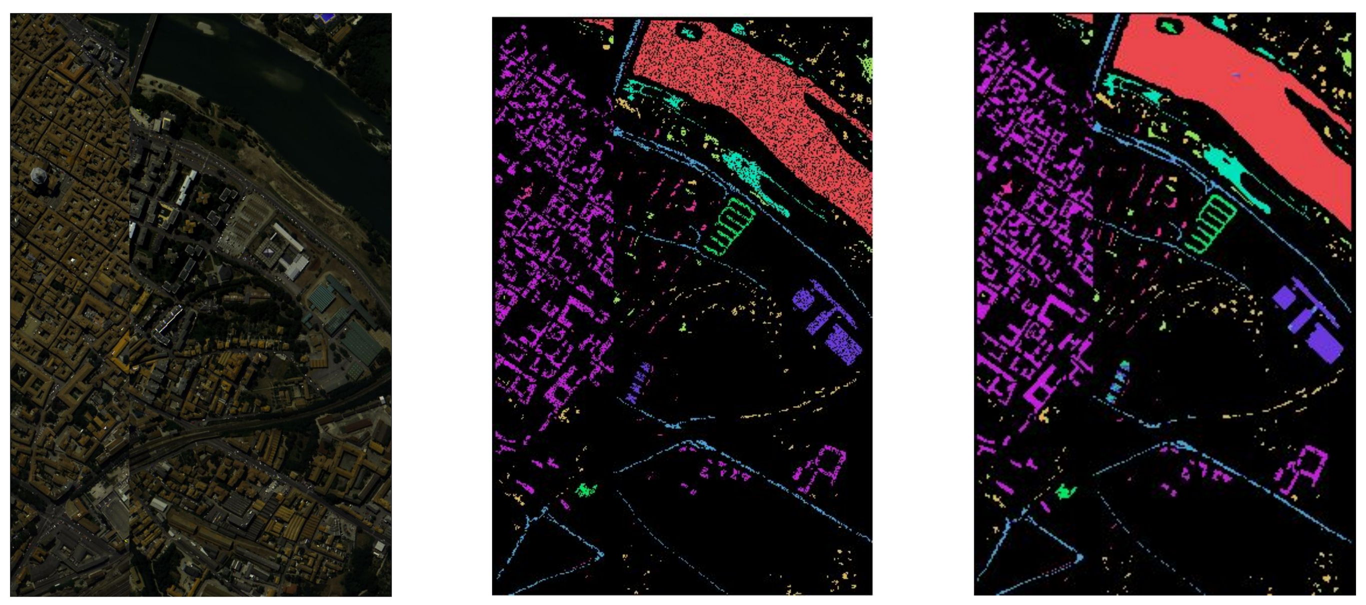
Image Classification:

Sort **MNIST Sign Language** and **SKLearn 'Labeled Faces in the Wild'** (LFW) data sets into categories

Methods

- Principal Component Analysis
- Implement Machine Learning Algorithms
  - Convolutional Neural Network (CNN)
  - Support Vector Machine (SVM)
  - K Nearest Neighbor (KNN)
  - Logistic Regression
  - Random Forest
- Analyze Results

Results: Image Segmentation



(a) (b) (c)

Figure 1. Pavia Centre data set with 9 classes. (a) Satellite image of Pavia, Italy (b) ground truth map (c) classified output from Hamida et al. 3D CNN, with 5% training data

Results: Image Segmentation

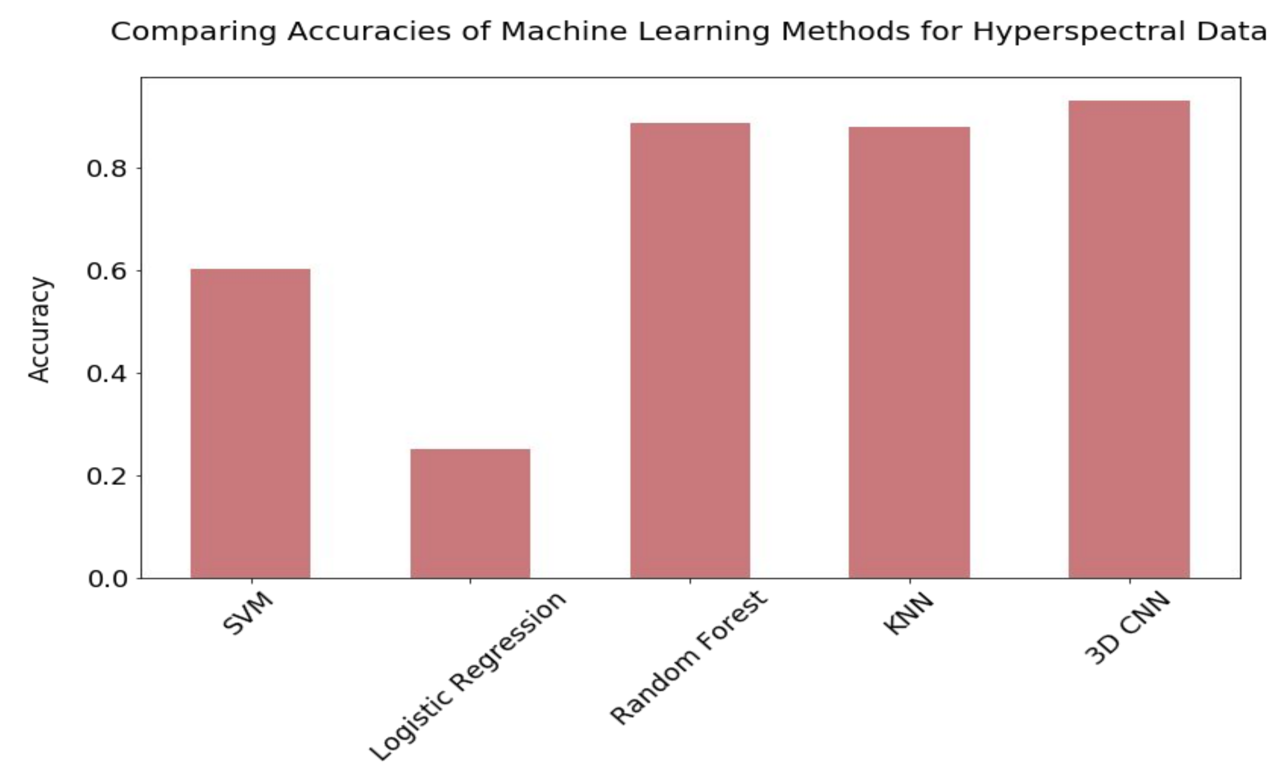


Figure 2. Accuracy of various machine learning algorithms for the hyperspectral data set, Pavia University, with 25% training data

Results: Image Classification

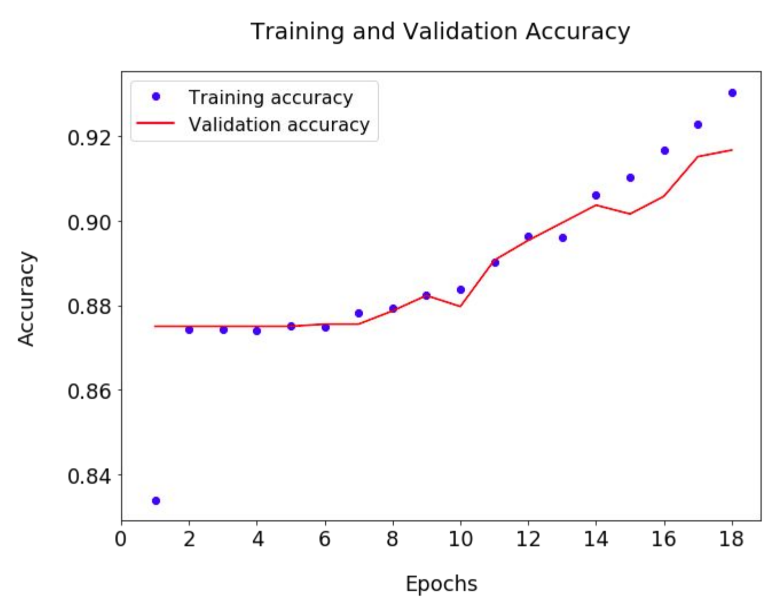


Figure 3. Training and validation accuracy versus number of epochs for the LFW data set using 15 layer CNN network with 25% training data

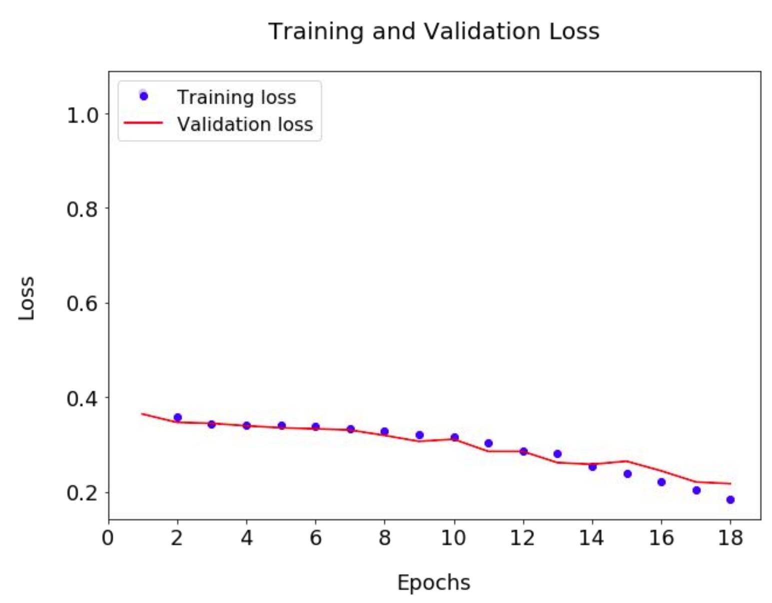


Figure 4. Training and validation loss versus number of epochs for the LFW data set using 15 layer CNN network with 25% training data

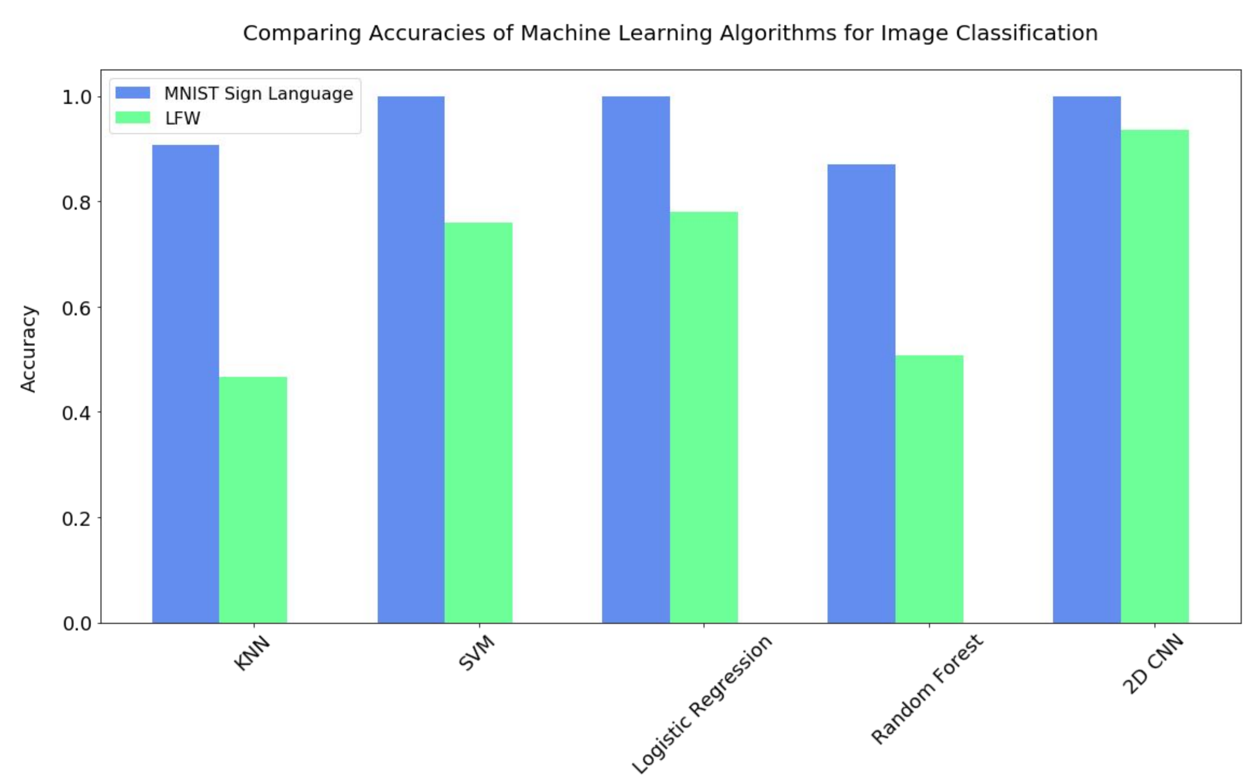


Figure 5. Accuracy of different machine learning algorithms for LFW and MNIST Sign Language data sets, with 25% training data

Results: Image Classification

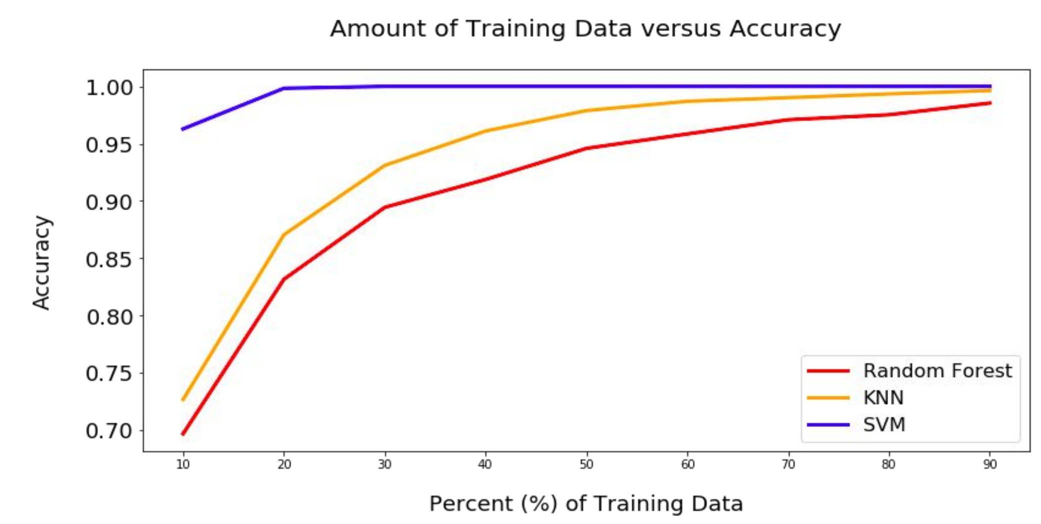


Figure 6. Accuracy versus amount of training data using SVM, KNN, and Random Forest models for MNIST Sign Language data set. Simpler methods such as KNN and Random Forest need higher amounts of training data to achieve higher accuracies.

Conclusion

Hyperspectral Image Segmentation:

- For hyperspectral images, **3D CNNs perform better than other traditional methods**. This is most likely due to 3D CNNs utilizing both spectral and spatial information.
- The Hamida et al. 3D CNN resulted in **97% accuracy with only 5% of training** data.

Image Classification:

- **15 layer CNN resulted in 93% accuracy** with no apparent overfitting, which was a common issue in CNNs with fewer layers.
- **2D CNN was the best performing model** for both data sets.
- MNIST dataset had significantly higher accuracies than LFW. However, facial images have more complex features compared to hand images which can possibly explain this difference.

References

We acknowledge the support of the MSU ACRES REU program and the NSF through grant OAC1560168.

[1] Gamba, Paolo. *Pavia University*. Telecommunications and Remote Sensing Laboratory. Pavia, Italy.
[2] Gamba, Paolo. *Pavia Centre*. Telecommunications and Remote Sensing Laboratory. Pavia, Italy.
[3] "Sign Language MNIST", *Kaggle*, 2017, https://www.kaggle.com/datamunge/sign-language-mnist.
[4] Pedregosa, Fabian, et al. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, vol. 12, 2011, pp. 2825-2830, http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf.
[5] Chollet, Francois, et al. *Keras,* 2015. https://keras.io.
[6] Hamida, Amina, et al. "3-D Deep Learning Approach for Remote Sensing Image Classification." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, 2018. pp. 4420-4434.