Literature Review Draft

Student: Daniel Hess - 21971897
Supervisors: Prof. Eun-Jung Holden, Dr Daniel Wedge

Introduction

With many methods being developed for the purposes of segmentation the field is becoming dense with options. There are many fields including medical, computer vision as well as autonomous vehicles which rely on segmentation systems to be fast a robust. In the medical field it allows researchers or technicians to receive more data from highly specialized imagery, while autonomous vehicles can accurately find road boundaries and vehicles.

This project will focus on data specific applications of segmentation for the mine pit data we're working on to accurately delineate detrital units from bedded units but will also look into public datasets that can be used that can be used as a benchmark for these techniques.

Many of the techniques focus on feature extraction from imagery and subsequent clustering of the features. This has the advantage that our feature space is only limited by our clustering algorithms, however can also become very large very quickly. The other major area in computer vision involves the use of Convolutional Neural Networks. These have the advantage that they can build deep convolutions which may suit our purposes well, but also are black box algorithms and may not be helpful from the perspective of insight. There are many other generalized models which get used in this field we will mention.

Finally, we can talk about metrics. How well the models perform with respect to our dataset will be crucial for this data focused project.

Feature Extraction

Gabor

Feature clustering is such an important area to segmentation from both a historical and performance perspective. The idea generally being that by using certain metrics, convolutions or statistics, we can extract information about the regions or pixels of an image, then based on this vector space, we can use conventional clustering techniques to split the pixels into different classes or regions.

Gabor filters are filters used in texture analysis, which in the 2D domain is comprised of a modulated Gaussian kernel. They are said to be similar to the human vision [6] and are particularly useful for texture and edge representation [texture classification using Gabor filters].

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma y'^2}{2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right)$$

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta$$

Xu et al [13] used gabor filters to not only extract features from grey-level images, but also from depth images which achieved good characterization of faces for the recognition process. The system benefitted from the inclusion of both intensity and depth features, compared to both individually.

Similarly, the use of log-Gabor filters have also been discussed. Nava et al [10] suggest that log- Gabor filters out perform traditional gabor filters.

Auto Encoders

Neural Networks (convolutional or otherwise) are such a large part of computer vision for a number of reasons. In this case we are interested in the use of auto-encoders as a form of feature extraction. Because the process of training auto-encoders includes a lower dimensional form of the data based on subsequent convolutions, using the features generated may be useful.

Yang et al [14] propose an auto-encoder based method for unsupervised defect detection. This method uses an auto-encoder for feature extraction which it clusters over to find surface defects. This method has the benefit that it is unsupervised however achieves great performance in finding surface defects.

Raja and rani [11] use a similar approach for tumor classification from MRI images. Using various other feature extraction methods, as well as features extracted from training an auto-encoder, the pair were able to achieve the highest accuracy from their method when compared to other existing systems. By incorporating the system with Bayesian fuzzy clustering, they were able to identify the positive regions in the MR images.

Clustering

With plenty of features extracted form the images, we need some way to select regions based on feature properties. A common method once features are extracted is to cluster these features. Provided we can effectively extract texture features from the images, clustering should provide a good framework for differentiating regions on texture.

K-means

One of the most popular clustering algorithms based on simplicity and fast computation. The kmeans algorithm can be summurised simply:
1. Take k points (call these the means) for each cluster.
2. Assign all observations to a cluster based on euclidean distance.
3. Calcualte the mean value for each given cluster and set this as the mean for that cluster.
4. Repeat steps 2-3 until observations do not change cluster.

The method has been shown to perform well for segmentation purposes based on feature extracted images. Lin et all [8] shows that given good features, decent segmentation can be achieved.

Fuzzy c-means

Similar to kmeans is fuzzy c-means. As opposed to hard clustering, this algorithm adopts soft clustering, meaning it considers the "probability" of a particular point being within a cluster. The algorithm is as follows:
1. Choose a number of clusters
2. Assign coefficients for observations to clusters randomly.
3. Compute the centroid for each cluster (i.e. the weighted mean for each point in the cluster)
4. Compute the coefficients for each cluster and observation.
5. Repeat 3-4 until the change in coefficients is less than some epsilon.

The main difference to kmeans being that we don't assign any specific observation to a cluster, and consider the "probability" of the point belonging to each cluster. This method allows for smoother transitions between clusters, and that centroids are weighted more heavily to points that are closer to it.

Yang et al [15] were able to segment normal and abnormal tissue in MR images by the use of fuzzy c-means clustering on intensity. Harikirian et al [4] achieved better results with the addition of other features for segmentation.

Finally Chuang et al [1] considered a fuzzy c-means algorithm that also considers spatial information. It is advantageous over normal FCM as it yield smroe homogenous regions and is also much less sensitive to noise.

DBSCAN

The last clustering algorithm we'll look at is DBSCAN. This algorithm takes a different approach to the other 2 as it considers the density of points within regions and will sometimes not assign observations to a cluster. The algorithm is as follows.
1. Find all points within epsilon size neighborhood of every observation.
2. Identify which points are core points by whether they have at least minPts neighbors.
3. Find which core points are connected to other core points and put them into their own cluster.
4. Assign each non-core point to a cluster if the cluster is within epsilon of a core point, otherwise assign it to noise.

The algorithm works well for arbitrarily shaped clusters and is robust to noise making it a well performing algorithm [7] however is very dependent on good hyperparameter selection.

Wang et al [12] proposed a method for automatic epsilon estimation which achieved better segmentation accuracy than methods that used different values of epsilon.

Convolutional Neural Networks

As mentioned earlier, neural networks have exceptional performance with most application. CNNs specifically have the benefit that working with images is done in the spatial domain and we can still get deep predictors.

While many CNNs have been trained to work as classifiers, work is still being done for segmentation. Long et al [9] managed to adapt architectures from previous state-of-the-art classifier systems into a segmentation system with good

performance. By discarding the final classifier layers, and using a 1x1 convolution, the classifier can now predict a score for each class over the image space, up-sampling wherever necessary.

A similar approach is done by Jégou et al [5]. The architecture for this method consists of a dense down-sampling path, reducing the convolutions in dimension, a bottleneck path to learn deep features followed by an up-sampling path with the final layer being dense, and finally a 1x1 convolution soft-max layer to provide class distribution. The distinction this makes with the previous system is that it uses a deep fully connected up-sampling path before segmentation.

A different approach taken by Fujieda et al [2] was to incorporate wavelets into the CNNs. While it follows a specific architecture, the idea is that it runs concurrently wavelet transforms alongside convolutional layers, with concatenation and shortcuts incorporated.

Metrics

Different metrics exist to measure goodness-of-fit, however different metrics solve different problems. Generally, in segmentation, there a few different metrics used to evaluate how well a method segments the data, which primarily fall into two categories: subjective and objective [3].

Objective evalutation is unsupervised, meaning it doesn't have labelled data to compare against as ground truth.

- F evaluation: Measures the average squared color error and works to regularise number of segments - has the benefit of not requiring hyper-parameters. Similarly, F' is another evaluation metric, to reduce over-segmentation further over F.

- Q evaluation: Another improvement over F', however improves bias towards under segmentation.

- E evalution: Compares the expected entropy of a segmented region to the actual measured entropy.

- Zeboudjs contrast: Based on the internal and external contrast contrast for the regions around a pixel.

Subjective evaluation considers a labelled ground-truth. Because we can consider segmented regions as classes, classification metrics work for evaluation on a pixel by pixel scale, however there are a few specific to segmentation.

- Probabilistic random index: Essentially a cluster-invariant evaluation of accuracy.

- Variation of Information: A combination of the entropy for both a cluster and the ground truth, as well as considering the mutual information between clusters. VOI measures the loss of information between clusters.

- Global Consistency Error: Consider a metric to evaluate the per pixel error of a segmentation. The GCE of a partition, is the scaled sum of error of all pixels in a region.

- Boundary Displacement Error: Consider the error on a boundary pixel, the error is defined as the distance between that boundary pixel, and the closest boundary pixel in the other ground-truth.

References

[1] CHUANG, K.-S., TZENG, H.-L., CHEN, S., WU, J., AND CHEN, T.-J. Fuzzy c-means clustering with spatial information for image segmentation. *computerized medical imaging and graphics 30*, 1 (2006), 9–15.

[2] FUJIEDA, S., TAKAYAMA, K., AND HACHISUKA, T. Wavelet convolutional neural networks. *arXiv preprint arXiv:1805.08620* (2018).

[3] GARCIA-LAMONT, F., CERVANTES, J., LÓPEZ, A., AND RODRIGUEZ, L. Segmentation of images by color features: A survey. *Neurocomputing 292* (2018), 1–27.

[4] HARIKIRAN, J., LAKSHMI, P., AND KUMAR, R. K. Multiple feature fuzzy c-means clustering algorithm for segmentation of microarray images. *International Journal of Electrical and Computer Engineering 5*, 5 (2015).

[5] JÉGOU, S., DROZDZAL, M., VAZQUEZ, D., ROMERO, A., AND BENGIO, Y. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (2017), pp. 11–19.

[6] JONES, J. P., AND PALMER, L. A. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of neurophysiology 58*, 6 (1987), 1233–1258.

[7] KURUMALLA, S., AND RAO, P. S. K-nearest neighbor based dbscan clustering algorithm for image segmentation. *Journal of Theoretical and Applied Information Technology 92*, 2 (2016), 395.

[8] LIN, W.-T., LIN, C.-H., WU, T.-H., AND CHAN, Y.-K. Image segmentation using the k-means algorithm for texture features. *World Academy of Science, Engineering and Technology 65* (2010), 612–615.

[9] LONG, J., SHELHAMER, E., AND DARRELL, T. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 3431–3440.

[10] NAVA, R., ESCALANTE-RAMÍREZ, B., AND CRISTÓBAL, G. A comparison study of gabor and log-gabor wavelets for texture segmentation. In *2011 7th International Symposium on Image and Signal Processing and Analysis (ISPA)* (2011), IEEE, pp. 189–194.

[11] SIVA RAJA, P., AND RANI, A. V. Brain tumor classification using a hybrid deep autoencoder with bayesian fuzzy clustering-based segmentation approach. *Biocybernetics and Biomedical Engineering 40*, 1 (2020).

[12] WANG, C., JI, M., WANG, J., WEN, W., LI, T., AND SUN, Y. An improved dbscan method for lidar data segmentation with automatic eps estimation. *Sensors 19*, 1 (2019), 172.

[13] XU, C., LI, S., TAN, T., AND QUAN, L. Automatic 3d face recognition from depth and intensity gabor features. *Pattern recognition 42*, 9 (2009), 1895–1905.

[14] YANG, H., CHEN, Y., SONG, K., AND YIN, Z. Multiscale feature-clustering-based fully convolutional autoencoder for fast accurate visual inspection of texture surface defects. *IEEE Transactions on Automation Science and Engineering 16*, 3 (2019), 1450–1467.

[15] YANG, M.-S., HU, Y.-J., LIN, K. C.-R., AND LIN, C. C.-L. Segmentation techniques for tissue differentiation in mri of ophthalmology using fuzzy clustering algorithms. *Magnetic Resonance Imaging 20*, 2 (2002), 173–179.