Multispectral Imaging for Improved Liquid Classification in Security Sensor Systems

Andrea Burns[a] and Waheed U. Bajwa[b]

[a]Center for Discrete Mathematics and Theoretical Computer Science (DIMACS)
[b]Department of Electrical and Computer Engineering
[a,b]Rutgers University-New Brunswick, Piscataway, NJ 08854

ABSTRACT

Multispectral imaging can be used as a multimodal source to increase prediction accuracy of many machine learning algorithms by introducing additional spectral bands in data samples. This paper introduces a newly curated Multispectral Liquid 12-band (MeL12) dataset, consisting of 12 classes: eleven liquids and an "empty container" class. Multispectral images in this dataset have been captured using the PCO Ultraviolet, Grasshopper3 12.3 MP Color USB3 Vision, Mil-Rugged-High Resolution Snapshot Short Wave Infrared 1280JS, FLIR Medium Wave Infrared A6750sc and FLIR Long Wave Infrared T650sc cameras. Each of the classes initially results in a $640 \times 480 \times 12$ data cube, where the 12×1 vector for each spectral pixel spans the spectral bands observed using the above-mentioned cameras and seven add-on bandpass optical filters. The usefulness of multispectral imaging in classification of liquids is demonstrated through the use of a support vector machine on MeL12 for classification of the 12 classes. The reported results are both encouraging and point to the need for additional work to improve liquid classification of harmless and dangerous liquids in high-risk environments, such as airports, concert halls, and political arenas, using multispectral imaging.

Keywords: Liquid classification; machine learning; multispectral imaging; support vector machine

1. INTRODUCTION

In recent years, multimodal data have made a huge impact on the field of machine learning, and they continue to do so because multimodal data better represent the way humans learn to perform tasks.[1,2] Despite the prominence of multimodality in our lives (e.g., understanding language through audio-visual information), there are challenges involved in using multimodal data in machine learning due to the difficulties of fusing various modes of different dimensions and data types.[3] One example of multimodal data, which have been successfully used in different types of classification tasks, is that of multispectral imaging (MSI). In MSI, multimodality corresponds to different spectral ranges; instead of using image data corresponding to the visual (typically, RGB) spectral range, a third dimension to the data is added in MSI, resulting in the same image observed over multiple spectral bands. In particular, whereas one RGB image may capture visible light in the $0.4 - 0.7$ μm wavelength range, a multispectral image spans a broader range of wavelengths from ultraviolet (UV) to infrared (IR) light, likely adding significantly more information to the data. The additional spectral bands in MSI can potentially help distinguish objects better, resulting in improved image classification performance compared to classification using only RGB or gray-scale images. Since MSI provides additional information outside of the traditional RGB spectral bands, it can be used to improve image classification within medical,[4,5] agriculture,[6,7] mineralogy, and military applications.[8] Specific applications where MSI and hyperspectral imaging (HSI) have provided advantages for image classification include vegetation and crop detection from remote-sensing devices,[9] and tumor detection in fMRI data.[10]

The application of interest in this paper is classification of liquids using MSI data, with potential usage in security sensor systems. Specifically, we provide initial evidence in this paper that MSI data can be used for improved classification of liquids. The results reported in the paper are based on training of a support vector machine (SVM) on a self-curated Multispectral Liquid 12-band (MeL12) dataset. Our results motivate the use of MSI in situations where liquids pose security concerns, allowing for improved screening at political events, concerts, and airports. The Transportation Security Administration (TSA) imposes strict limitations on what

Class #	Class Name	Total Number of "Spectral" Pixels
1	Coca Cola	53,998
2	Minute Maid Cranberry Apple Raspberry Juice	52,800
3	Empty Container	51,084
4	Seagram's Ginger Ale	52,800
5	Hydrogen Peroxide	53,067
6	Lemonade Vitamin Water	53,200
7	Milk	50,625
8	Minute Maid Orange Juice	53,198
9	Rubbing Alcohol	52,400
10	Seagram's Seltzer Water	52,400
11	Soy Milk	52,668
12	Water (tap)	52,668

Table 1. Summary of MeL12 dataset in terms of the 12 liquid classes

liquids are allowed in carry-on luggage, as well as 3.4 oz size restrictions on the allowed liquids. With the accuracy achieved in this work, the TSA may no longer have to maintain the severity of these restrictions. Liquids could be automatically classified in clear bottles, or at the very least categorized into benign versus hazardous groups for improved travel security and flyer experience. Additional data curation of dangerous liquids would strengthen these results also.

We conclude this introduction by noting that sparse representations and deep learning models are now routinely used in machine learning for state-of-the-art object recognition.[11] However, these approaches typically require large volumes of training data to avoid overfitting.[12] In contrast, SVMs typically exhibit good performance in the presence of limited training data.[13] Because of this reason, and since the purpose of this paper is mainly to provide a proof-of-concept for the use of MSI in image classification, we limit ourselves to SVM-based classification of liquids using MeL12 throughout the rest of this paper.

2. RESEARCH METHODOLOGY

2.1 Acquisition of Dataset

The Multispectral Liquid 12-band (MeL12) dataset comprises 12 classes of liquids, which are listed in Table 1. In addition to differences in liquid properties such as density and viscosity, some of the classes also vary from each other in terms of color, opaqueness, and carbonation. Note, however, that there are several clear liquids and cream/white liquids included in MeL12, which are meant to demonstrate the usefulness of MSI in differentiating liquids of same or similar color. All images in MeL12 dataset were taken with an identical set up: a background poster was placed with alignment boxes in each corner (similar in spirit to a QR code) to ensure uniform viewing frames and straightforward post-acquisition alignment procedures. The same Poland Spring 16 oz water bottle was used for holding each liquid class, which was placed on top of a small stand to create a better imaging environment. Markers on the cap of the bottle were used to help align the bottle between the use of cameras and the switching of liquids. Of the 12 spectral bands comprising each length-12 pixel in MeL12, five bands were observed directly using five different cameras, while the remaining seven bands were observed through the use of seven bandpass filters added to the cameras to further discretize the spectral ranges of individual cameras. Detailed specifications of these five cameras and seven bandpass filters are provided in Table 2. Note that the filter model listed in Table 2 also provides the median wavelength and the bandwidth of the filter (in nanometers); e.g., the filter FL730-10 has a median wavelength of 730 nm and a bandwidth of 10 nm, therefore approximately spanning the spectral wavelength of $0.725 - 0.735$ μm. It is also worth mentioning here that the 12 spectral bands being captured in MeL12 are not strictly disjoint; in particular, the spectral bands being directly observed by the short wave infrared (SWIR) and medium wave infrared (MWIR) cameras overlap with the spectral bands captured by some of the filters. The effects of these overlaps and discretization of the spectral bands on classification accuracy are explored later in the paper. We refer the reader to Fig. 1 for a graphical representation of how different cameras and filters capture the 12 bands within MeL12 dataset.

Camera/Filter	Model and Lens	Spectral Range (Wavelength)
UV Sensitive Camera	PCO pco.ultraviolet Edmund Optics Lens 57-542 f/2.8 - f/16, 25mm	$0.2 - 0.7$ μm
Visible Light (VIS) Reference Camera	Point Grey Research Grasshopper3 GS3-U3-123S6C-C Schneider Optics Xenon Topaz f/2.0, 38mm	$0.4 - 0.7$ μm
High-Resolution Short Wave Infrared (SWIR) Camera	UTC Aerospace GA1280JS UTC Aerospace Solo 50mm	$0.9 - 1.7$ μm
Bandpass Filter	Thorlabs FL730-10	$0.725 - 0.735$ μm
Bandpass Filter	Thorlabs FBH800-10	$0.795 - 0.805$ μm
Bandpass Filter	Thorlabs FBH850-10	$0.845 - 0.855$ μm
Bandpass Filter	Thorlabs FL900-10	$0.895 - 0.905$ μm
Bandpass Filter	Thorlabs FLH1030-10	$1.025 - 1.035$ μm
Bandpass Filter	Thorlabs FBH1200-10	$1.195 - 1.205$ μm
Bandpass Filter	Thorlabs FBH1550-12	$1.544 - 1.556$ μm
Medium Wave Infrared (MWIR) Sensitive Camera	FLIR A6750sc f/2.5 Lens, 50mm	$1 - 5$ μm
Long Wave Infrared (LWIR) Sensitive Camera	FLIR T650sc, f/1.0 Integrated Lens, 18mm	$7.5 - 13$ μm

Table 2. Technical specification of the equipment used to acquire MeL12 dataset

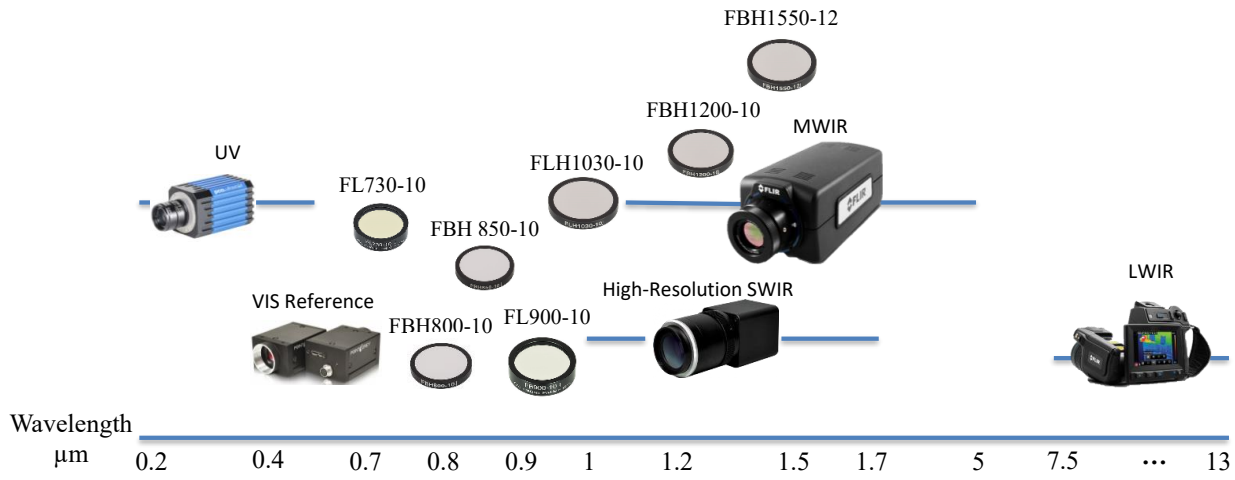

Figure 1. The MeL12 dataset is captured using five cameras (UV, VIS, SWIR, MWIR, and LWIR) and seven bandpass filters added to some of these cameras. The relationship between the cameras, filters, and the spectral bands captured in MeL12 is being summarized in this figure.

2.2 Post Processing of Dataset

Due to differences in various specifications of the five cameras, there were several variations among the starting set of captured images. Therefore, a number of post-processing steps were carried out to bring uniformity across different images so that they could be combined into a single multispectral data cube for each liquid class. First, any RGB images were converted to grayscale images; in addition, bit depth of all images was adjusted to 8-bit so that all images had pixel values in the $0-255$ range. Next, while different cameras took images with varying pixel resolutions, all images were downsampled to the lowest resolution (corresponding to the LWIR camera) of 640×480 pixels. Further, since the five cameras resulted in different fields of view due to factors such as focal lengths of the lenses, this severely affected the viewing frame for each image. This issue was partly addressed by physically adjusting the imaging environment between uses of different cameras. Despite this, however, focusing on the bottle during acquisition caused the cameras to occasionally alter their fields of view, which disrupted the

alignment between images. We addressed this issue by registering the images of each class in different spectral bands in reference to the VIS image of that class (after correcting for bit depth and pixel resolution).

The aforementioned post-processing steps gave rise to a three-dimensional data "cube" of dimension $640 \times 480 \times 12$ for each liquid class, resulting in a total of 12 data cubes. Afterward, we cropped these data cubes to retain only spectral pixels corresponding to the bottle, thereby excluding the background from MeL12. We then labeled these pixels from $1-12$ according to their corresponding class and finally "vectorized" the 12-dimensional pixels into a $630,908 \times 12$ data cube for the entire dataset. In other words, a single sample in the (post-processed) MeL12 dataset is a 12-dimensional pixel, spanning the 12 spectral bands, and its corresponding label.

2.3 Classifier Training

In order to evaluate the usefulness of multispectral data in classifying liquids within containers, we trained *one-againt-one* linear SVM classifiers on MeL12 dataset. The data was split into training and test sets using randomized cross validation, with the reported results averaged over five trials and corresponding to 90% and 10% split among training set and test set, respectively. Before training, data in MeL12 was shifted and scaled to have zero mean and unit variance, respectively. The final decision for liquid classification was made with a majority vote among the classifiers.

3. NUMERICAL RESULTS

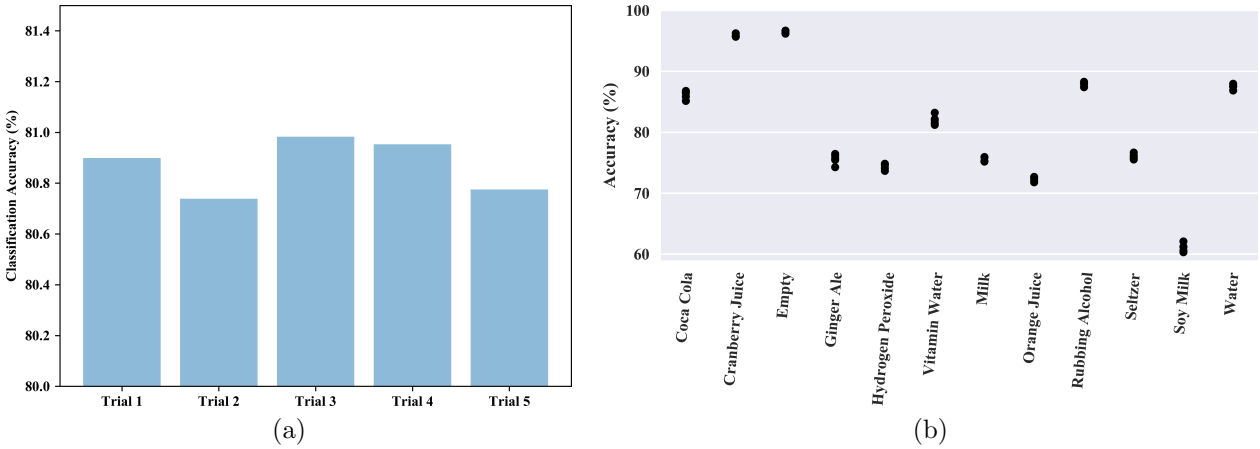

Figure 2. Classification accuracy of linear binary SVMs on all 12 bands of MeL12 with a final decision rule of majority voting. (a) Classification accuracy averaged over the 12 classes as a function of different trials. (b) Classification accuracy of individual liquid classes for the five trials.

We first focus on the case of liquid classification using all 12 bands of MeL12. The classification accuracy of the trained SVMs in this case is shown in Fig. 2, with Fig. 2(a) showing the classification accuracy averaged over all 12 classes as a function of different trials and Fig. 2(b) displaying the classification accuracy for each individual class for each of the randomized trials. It can be seen from Fig. 2(a) that even classifiers as simple as linear SVMs lead to average classification accuracy of around 80.8%. We consider this classification accuracy to be encouraging for a collection of 12 liquids that have significant overlaps in physical properties such as color. We also see from Fig. 2(b) that the classification accuracy of most classes varied from 70% to 90%. However, cranberry juice achieved accuracy in excess of 90%, while soy milk achieved accuracy closer to 60%. In order to better understand the low classification accuracy of soy milk, we also provide a confusion matrix representation of per-class classification accuracy in graphical and numerical forms in Fig. 3 and Table 3, respectively; here, the rows correspond to the true class labels and the columns represent the predicted labels. It can be seen from the confusion matrix that soy milk is most frequently misclassified as orange juice, milk, Coca Cola, and cranberry juice. Three other significant cases of misclassification occurred between hydrogen peroxide and rubbing alcohol, ginger ale and seltzer, and vitamin water and orange juice. These errors suggest the need for further investigation of more refined machine learning algorithms for liquid classification using MeL12.

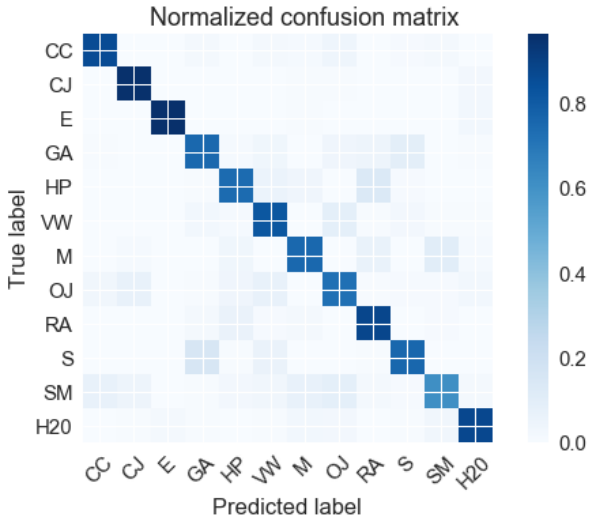

Figure 3. A graphical representation of the normalized confusion matrix for the 12 classes in MeL12 dataset, averaged over five randomized trials. The abbreviations representing different classes are: CC = Coca Cola, CJ = Cranberry Juice, E = Empty Container, GA = Ginger Ale, HP = Hydrogen Peroxide, VW = Vitamin Water, M = Milk, OJ = Orange Juice, RA = Rubbing Alcohol, S = Seltzer, SM = Soymilk, and H20 = Water.

	CC	CJ	E	GA	HP	VM	M	OJ	RA	S	SM	H20
CC	**86.19**	0	.06	2.23	.24	2.11	1.5	4.38	.12	1.06	2.12	0
CJ	0	**95.86**	.07	0	0	0	.64	.7	0	.1	.25	2.38
E	0	0	**96.49**	0	0	0	.52	0	0	0	.06	2.93
GA	.41	0	0	**75.61**	1.2	3.47	.1	4.12	4.63	9.43	.3	.72
HP	.02	0	0	.43	**74.39**	5.97	3.92	.32	13.56	1.09	.28	.01
VM	.03	0	0	2.99	1.5	**82.03**	.02	8.91	.71	2.48	.63	.7
M	.18	1.3	.37	.05	3.68	.06	**75.53**	.05	6.74	.22	10.41	1.42
OJ	3.32	6.79	.17	.7	3.92	7.47	.04	**72.4**	.7	1.05	.78	2.66
RA	0	.01	0	1.62	6.26	.92	1.58	.09	**87.85**	.35	1.06	.25
S	0	.08	.07	15.58	.45	6.23	.13	.34	.62	**76.01**	.25	.24
SM	6.93	4.64	.27	.42	2.3	3.33	7.11	8.85	2.09	.76	**61.09**	2.21
H20	0	.6	1.67	.49	0	.42	2.88	2.57	.14	.4	3.3	**87.5**

Table 3. A numerical representation (in %) of the confusion matrix for the 12 classes in MeL12 dataset, averaged over five randomized trials.

In addition to performing classification using all 12 bands of MeL12 dataset, we also investigated the role of the number and spectral ranges of bands in accuracy of liquid classification. We noticed in our experiments that, in general, the classification accuracy decreased with a decrease in the number of spectral bands; however, the drop in accuracy greatly depended on the bands that were removed from the original MeL12 dataset. We report these results in Table 4, which lists the average accuracy as a function of the number of bands as well as the best and worst classification accuracy as different spectral ranges are removed from the dataset. It can be seen from these results that there is a significant variation in classification accuracy as a function of the spectral ranges of different bands. In particular, when using 11, 10, 9, or 8 bands out of the total 12 for classification, not only is the average classification accuracy lower, but also there is a variation of $3-20\%$ approximately in accuracy as different bands are removed. In addition, while not listed in Table 4, accuracy of a third of the classes (vitamin water, milk, orange juice, and seltzer water) reduced to between 40% and 50% with the removal of as few as three bands (i.e., using 9 bands for classification). When using 11 bands for classification, we noted that removal of the UV band impacted the classification accuracy the most, resulting in final accuracy figure of 75.3%. When using five bands for classification, we limited ourselves to the spectral ranges natively covered by the five cameras used in our study (i.e., without the add-on bandpass filters). In this case, we observed

Number of Bands	Accuracy Results (averaged over the 12 classes)
12 bands	Average over 5 runs: 80.87%
11 bands	Average over all possible uses of 11 bands: 78.43% Best: 80.62% (without FBH850-10 filter) Worst: 75.3% (without UV spectral band)
10 bands	Average across different spectral ranges: 73.5% Best: 79.49% (without FL730-10 and FBH800-10 filters) Worst: 65.24% (without Visible Light and MWIR spectral bands)
9 bands	Average across different spectral ranges: 65.6% Best: 78.70% (without FL730-10, FBH800-10, and FBH850-10 filter) Worst: 55.5% (without UV, Visible Light, and MWIR spectral bands)
8 bands	Average across different spectral ranges: 64.77% Best: 77.13% (without FL730-10, FBH800-10, FBH850-10, & FL900-10 filters) Worst: 55.73% (without UV, SWIR, MWIR & LWIR spectral bands)
7 bands	Average using only filter bands: 41.42%
6 bands	Average across different spectral ranges: 58.87% Best: 60.26% (using the 2^{nd} half of spectral bands) Worst: 57.48% (using the 1^{st} half of spectral bands)
5 bands	Average using bands without filters: 67.15%
1 band	Average over 5 runs using only visible light: 24.11%

Table 4. Average, best, and worst classification accuracy for the 12 liquid classes as a function of the number and spectral ranges of different bands retained within the MeL12 dataset.

average classification accuracy of 67.15%, which is a drop of more than 10% from 12 bands. This points to the usefulness of discretization of the spectral ranges spanned by the five cameras using bandpass filters. Finally, we also performed classification using only the visible spectral band; the average classification accuracy in this case dropped to 24.11%, with a majority of the classes achieving an accuracy of 10% or lower. This again points to the usefulness of multispectral imaging in classification of liquids.

We conclude this section with a final set of observations. First, we noticed that the "water" class needs the MWIR band to maintain classification accuracy over 80%; in the absence of this band, the accuracy of this class dropped to around 30%. This is not surprising since the MWIR spectral range of $1-5$ μm includes the water absorption bands. In addition, classification accuracy of the "ginger ale" class dropped to the 50% range without the UV and LWIR bands, while it dropped to the high 40% range in the absence of the UV and MWIR bands. We also noticed that the "seltzer" and "soymilk" classes are the most affected by removal of any band from the dataset. Further, we noticed that classification accuracy of the "vitamin water" class dropped down to the 50% range or less with 10 bands or lower, with most misclassification errors due to it being confused with orange juice. In contrast, however, the "orange juice" class maintained its classification accuracy in the 80% range with 10 bands. Finally, when using six bands, accuracy of the "rubbing alcohol" class reduced significantly in the absence of the second half of the spectral bands, with the first six bands resulting in 24% accuracy and the second six bands giving rise to 82% accuracy.

4. CONCLUSION

In this paper, we focused on the problem of liquid classification using multispectral imaging (MSI). This involved curating a 12-band dataset, termed MeL12, comprising 12 classes of liquids using a collection of five cameras and seven add-on bandpass filters. In order to evaluate the usefulness of MSI in liquid classification, we trained binary linear SVMs on the MeL12 dataset and demonstrated an average classification accuracy of around 80.5%, which is significantly higher than that achievable using RGB images alone. We also investigated the role of the number and spectral ranges of different bands on the final classification accuracy, and highlighted the usefulness of discretization of spectral bands in high-bandwidth cameras using bandpass filters. The outcomes of this study suggest that as few as 12 spectral bands can be used to significantly enhance the accuracy of classification tasks involving multispectral data. It stands to reason that one should be able to improve further on these results using additional spectral bands and/or larger training datasets.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of National Science Foundation through grant CCF-1559855 and Army Research Office through DURIP Award W911NF-16-1-0209. We are also thankful to Prof. Mark Pierce for his help during collection of the MeL12 dataset. Andrea Burns completed this work as part of her 2017 REU summer project at DIMACS, Rutgers University-New Brunswick.

REFERENCES

[1] Baltrusaitis, T., Ahuja, C., and Morency, L.-P., "Multimodal machine learning: A survey and taxonomy," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PP**(99), 1 (2018).

[2] Ngiam, J., Khosla, A., Kim1, M., Nam, J., Lee, H., and Ng, A. Y., "Multimodal deep learning," *Proc. of the 28th Intl. Conf on Machine Learning (ICML-11)* , 689–696 (2011).

[3] Lahat, D., Adali, T., and Jutten, C., "Multimodal data fusion: An overview of methods, challenges and prospects," *Proc. of the IEEE* **103**, 1449–1477 (09 2015).

[4] Siddiqi, A. M., Li, H., Faruque, F., Williams, W., Lai, K., Hughson, M., Bigler, S., Beach, J., and Johnson, W., "Use of hyperspectral imaging to distinguish normal, precancerous, and cancerous cells," *Cancer Cytopathology* **114**(1), 13–21 (2008).

[5] Vo-Dinh, T., Kasili, P. M., and Cullum, B. M., "Multispectral imaging for medical diagnostics," *Proc. of SPIE* **4615**, 13–19 (2002).

[6] Qin, J., Chao, K., S.Kim, M., Lu, R., and F.Burks, T., "Hyperspectral and multispectral imaging for evaluating food safety and quality," *Journal of Food Engineering* **118**(2), 157–171 (2013).

[7] Lamb, D. W., "The use of qualitative airborne multispectral imaging for managing agricultural crops - a case study in south-eastern Australia," *Australian Journal of Experimental Agriculture* (40), 725–738 (2000).

[8] Briottet, X., Boucher, Y. G., Dimmeler, A., Malaplate, A., Cini, A., Diani, M., Bekman, H., Schwering, P., Skauli, T., Kasen, I., Renhorn, I., Klasen, L., Gilmore, M., and Oxford, D., "Military applications of hyperspectral imagery," *Proc. of SPIE* **6239**(62390B), 1–8 (2006).

[9] Baumgardner, M. F., Biehl, L. L., and Landgrebe, D. A., "220 Band AVIRIS hyperspectral image data set: June 12, 1992 Indian Pine test site 3," *Purdue University Research Repository* (2015).

[10] Szilgyi, L., Lefkovits, L., and Beny, B., "Automatic brain tumor segmentation in multispectral MRI volumes using a fuzzy c-means cascade algorithm," in [*2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*], 285–291 (Aug 2015).

[11] Chen, Y., Nasrabadi, N. M., and Tran, T. D., "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Transactions on Geoscience and Remote Sensing* **49**(10), 3973 – 3985 (2011).

[12] Cawley, G. C. and Talbot, N. L. C., "On over-fitting in model selection and subsequent selection bias in performance evaluation," *Journal of Machine Learning Research* **11**, 2079–2107 (2010).

[13] Melgani, F. and Bruzzone, L., "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing* **42**(8), 1778 – 1790 (2004).