A Deep Semantic Mobile Application for Thyroid Cytopathology

Edward Kim[1], Miguel Cortre-Real[1], Zubair Baloch[2]

[1]Department of Computing Sciences, Villanova University, Villanova, PA
[2]Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA

ABSTRACT

Cytopathology is the study of disease at the cellular level and often used as a screening tool for cancer. Thyroid cytopathology is a branch of pathology that studies the diagnosis of thyroid lesions and diseases. A pathologist views cell images that may have high visual variance due to different anatomical structures and pathological characteristics. To assist the physician with identifying and searching through images, we propose a deep semantic mobile application. Our work augments recent advances in the digitization of pathology and machine learning techniques, where there are transformative opportunities for computers to assist pathologists. Our system uses a custom thyroid ontology that can be augmented with multimedia metadata extracted from images using deep machine learning techniques. We describe the utilization of a particular methodology, deep convolutional neural networks, to the application of cytopathology classification. Our method is able to leverage networks that have been trained on millions of generic images, to medical scenarios where only hundreds or thousands of images exist. We demonstrate the benefits of our framework through both quantitative and qualitative results.

Keywords: cytology, thyroid, semantic web, mobile, image retrieval, ontology, deep learning, machine learning, convolutional neural networks

1. INTRODUCTION

A very large variety of thyroid lesions exist, ranging from the benign follicular nodule to the malignant neoplasms. With thyroid cancer on the rise[1] and guidelines for treatment of thyroid disease evolving,[2] it is important that physicians have access to resources and tools that can assist them in the process of managing thyroid disease. Currently, management and diagnosis utilizes thyroid fine needle aspiration (FNA) in conjunction with ultrasonographic imaging. The interpretation of FNA samples can be challenging due to the visually diverse preparation, which include smears, cell blocks, and stains (Papanicolaou, Romanowsky-Type, H&E, etc.), that can contain any number of possible anatomical and pathological characteristics. Thus, we propose a semantic mobile application that can serve as a reference or can assist with the management and interpretation of digitized FNA images.

Our goals for this application are two-fold. The first and primary goal is to create a mobile reference and educational application of thyroid cytopathology. The reference will be free for all users and available on the major mobile operating systems (Android, iOS) as well as standard web browsers. Our application has several benefits over static references and atlases.[3–5] Because the application data is centralized, it can be easily updated and changed if management or diagnostic procedures are revised. Another benefit over static resources is the wide accessibility of the application, both in terms of cost and availability at the point of care, which are both very important for low resource regions of the world.

Our second goal is to create a novel mobile application with semantic capabilities. Semantic data has the ability to share and link the information created to other existing resources. This includes linking to larger systems including the unified medical language system (UMLS),[6] NCI Thesaurus,[7] and Wikipedia via DBpedia.[8] Another benefit is the ability of our mobile application to act "intelligently" utilizing the underlying ontology developed for the system. This is particularly useful in our application where we have the ability to search for similar images or search for similar pathologies. Because relationships are defined or discovered as a result of the semantic representation, we are able to perform complex image and data queries. Our semantic layer adds an additional layer of knowledge over existing online resources and atlases,[9, 10] which only have search capabilities

for figures metadata and keywords. These existing interactive systems do not support direct image based searches nor provide a workflow for thyroid disease diagnosis.

The semantic labelings in our application are manually obtained; however, we explore the possibility of augmenting our application with recent advances in machine learning and computer assisted classification. Of particular benefit is the use of pretrained convolutional neural networks (CNNs)[11–14] for feature extraction and classification. These networks have been trained by massively multicore GPUs on millions of generic images. In the following sections, we will describe how we can utilize these CNNs for much smaller scale classification tasks where we do not have millions of images available from which to train, while simultaneously maintaining the benefits of the large training data corpus. Ultimately, we will show that the utilization of these methods can translate to clinical relevance and semantic labeling with the interpretation of digitized FNA images.

2. BACKGROUND

2.1 Semantic Technologies

Our semantic capability extends from our underlying pathology ontology. Ontologies define a vocabulary, concepts, and their relationships within a domain of use. Establishing a vocabulary is particularly helpful in multimedia annotation and retrieval as it provides a standard language for annotation; however, the integration of multimedia and metadata is performed in different ways and is still insufficiently explored.[15, 16] Semantic Web languages including Resource Description Framework (RDF), Web Ontology Language (OWL) are becoming more popular and new tools are emerging supporting knowledge bases on mobile platforms.[17] Tolksdorf et al.[18] developed a related platform that engineered an ontology for pathology using RDF/OWL. Our system further builds on this work by incorporating additional capabilities of automatic image property classification using machine learning techniques. Moreover, our system is engineered to be compatible for cross-platform mobile frameworks. There are other mobile medical apps and atlases, *Pocket Atlas of Anatomy*, *Pathology Dictionary*, *Anatomy - Cytology*; however, most are basic applications, and to our knowledge, none of these mobile applications utilize semantic technology to represent their data.

2.2 Computer Assisted Digital Pathology

Traditional pathology is the examination of stained tissue samples under a microscope. In the traditional scenario, the pathologist is limited to a single magnification viewing, manual slide analysis, and cumbersome archival and retrieval. The results of manual inspection by expert pathologist can also suffer from considerable inter and intra-reader variability.[19] The digitization of pathology aims to address many of these issues but has a critical hurdle, the amount of data generated is enormous. In contrast to radiological modalities such as a CT scan (that may consist of a 3 dimensional volume of size of ~134 million voxels), a single biopsy digitized at 40x resolution consists of 15,000 x 15,000 elements, or (~225 million pixels). A patient may have 12-20 slides resulting in 2.5-4 billion pixels of data.[20] Additional challenges for automated methods include the large variability present including cell types, cell densities, stains, and magnification levels. Analysis of cytology imagery obtained from biopsies is relatively easier (but still difficult) for computational algorithms due to the presence of isolated cells and the absence of more complicated anatomy. Histological images are more challenging due to the underlying tissue architecture, but have the benefit of providing a more comprehensive view of a disease.[20]

2.3 Cytohistology Features

For example, let us consider the FNA image samples in Figure 1. We show two benign lesions (a)(b) and two malignant tumors (c)(d) using the Papanicolaou stain. The benign tumors exhibit the following features,[3]

- **Cell Patterns** : Predominant macrofollicular pattern of follicular cells in flat honeycomb sheets

- **Cells** : Scant delicate cytoplasm

- **Nucleus** : Small, round nucleus with granular dark chromatin, regular nuclear membranes

- **Background** : Abundant watery or thick colloid admixed with variable amounts of blood and cystic components

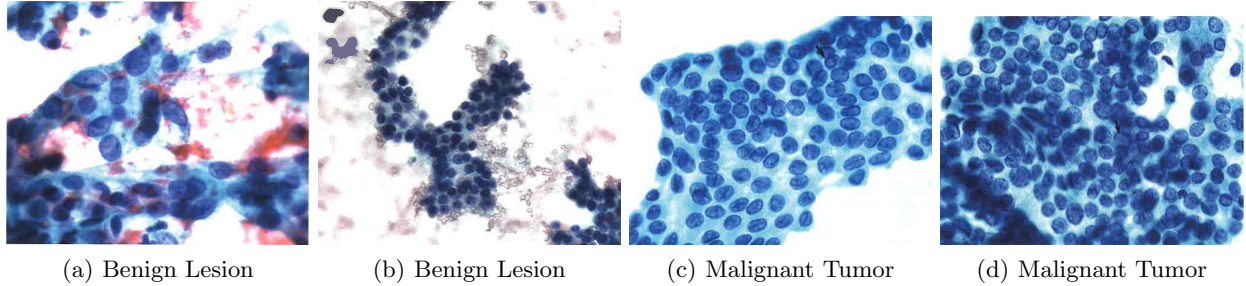| (a) Benign Lesion | (b) Benign Lesion | (c) Malignant Tumor | (d) Malignant Tumor |

Figure 1. (a) Benign thyroid nodule with reactive nuclear changes. (b) Benign nodule with follicular cells mostly in flat sheets. (c) Papillary carcinoma cells showing prominent nuclear elongation, chromatin clearing. (d) Papillary carcinoma demonstrating a cellular whorl, intranuclear grooves. All images are Papanicolaou-stained and obtained from Baloch et al.[3]

In contrast, the malignant tumors exhibit the following features.

- **Cell Patterns** : Picket fenced in papillae, honeycomb in syncytial fragments with evident cellular border

- **Cells** : Larger than normal, wide variation in size and shape

- **Nucleus** : Large, oval-shaped, regular to irregular in shape and overlapping. Grooves and pseudo-inclusions, clear and pale nuclear chromatin

- **Background** : Scant colloid, cystic background in cystic variant, branching or readily evident fibrovascular cores papillary structures.

Our semantic application is designed to provide these types of annotations and link them to a larger knowledge base. Our machine learning capabilities are designed to take these annotations for training classification algorithms that could be used for automated pathology image analysis.

3. METHOD

Our application consists of several components. The medical content is created and stored in an ontology. We built a machine learning classifier that is able to translate image characteristics and features into RDF/OWL object properties. In order to access the ontology through a mobile platform, we utilize several open source libraries and resources, including PhoneGap,[21] Apache Jena (and Jena Fuseki),[22] and JQuery Mobile.[23] Each of these components is described in more detail in the following sections.

3.1 Ontology

For the development of our thyroid ontology, we created a knowledge base around the cytomorphology of benign nodular goiters, oncocytic lesions, and malignant papillary and medular carcinoma. The ontology was created using the Protégé system and stored on a central server using Apache Fuseki. We are able to query the ontology using the SPARQL query language through a RESTful web interface.

3.1.1 Image Properties

Our ontology is focused on describing the visual characteristics of disease and the presence or lack of presence of certain anatomic characteristics. Our ontology can describe the magnification of the image (*hasMagnification*), the stain of the image (*hasStain*), the disease or lesion type present in the image (*exampleOfDisease*), the cell properties (*cellsDemonstrate*), as well as the cell morphology (*hasArtifact, hasColor, hasShape, hasSize, hasConsistency, hasArrangement*).

Each class in our ontology is linked to the nearest associated term in the NCI thesaurus and DBpedia. For example, if we consider the resource, "Hürthle Cell" , we find the following definitions, (DBpedia dbo:abstract): "A Hürthle cell is a cell in the thyroid that is often associated with Hashimoto's thyroiditis as well as follicular

thyroid cancer" and (NCI Thesaurus ncicp:ComplexDefinition): "A morphologic finding indicating transformation of follicular cells to large cells with abundant eosinophilic and granular cytoplasm (Hürthle cells) in the thyroid gland." These definitions are linked so that a user can see definitions of these classes within our mobile application. There are two automatically populated image properties added to the images, the *hasComputedStain* and *hasComputedStainConfidence*, that use our machine learning techniques for classification of images. In the following sections, we describe how these properties were computed.

3.2 Feature Extraction and Classification

The first step to automatic classification involves the extraction of discriminatory image features. Some features that one could collect have clear pathological value such as size, area, shape (elliptical features); however, the extracted features that attempt to describe the color and texture of regions (which will be used in this application) are human engineered features, (histogram of oriented gradients,[24] co-occurance gray level matricies,[25] chromatin density[26]) and not derived from the data itself. The classifier of this data must be able to deal with large, dense datasets.[20] Ensemble schemes,[27] support vector machines,[28] and Adaboost[29] classifiers have been popular choices in automatic cytopathology classification.

As we briefly mentioned, recognition in this domain has relied on hand-designed features, but these features only capture low-level information. The hand engineered features are not created directly, but rather are parameter tuned, from the training and testing data. These classifiers are also shallow, meaning they have one or zero feature transformation layers between the data and classification (SVM, Gaussian mixture models, etc). These models work well for well-constrained problems, but are limited when looking at real world image and visual representations. Our work utilizes deep (multiple transformation layers) convolutional neural networks which capture the low-level information, as well as mid-level and high level representations of objects generated directly from the data itself.[12, 30]

3.3 Details of our Image Feature Descriptors

Color Feature - We use a color image feature, PLAB.[31] Pixel colors are transformed into the perceptually uniform $L^*a^*b^*$ color space and for each channel (L^*, a^*, or b^*) of the color space, we extract 3 pyramid levels, with a 16 bin histogram from each region. A pyramid is constructed by splitting the image into rectangular regions, increasing the number of regions at each level. Thus, a single channel histogram consists of 336 bins, and our complete PLAB descriptor consists of 1008 bins.

Edge Feature -The edge is represented as a pyramid histogram of oriented gradients, or PHOG feature.[32] To extract the PHOG descriptors from an image or image region, we first compute the gradient response using a sobel edge filter. If we use an 8 bin orientation histogram over 4 levels, the total vector size of our PHOG descriptor for each image is 680 bins.

Texture Feature - The texture information is captured in the PLBP[33] features, or a pyramid of local binary patterns from the same pyramid regions. Each level consists of 58 bins, for a total of 1218 bins over the entire image.

CNN codes - CNN codes image features extracted from a convolutional neural network output on a hidden layer. Our approach, which is a typical approach, is to remove the last fully connected classification layer and use the 4096 dimensional vector activation outputs as the image feature. These activations are thresholded at zero by a ReLU activation function and are obtained for both training and testing in a separate classifier (like a Linear SVM).

3.4 Details of our Machine Learning Methods

Random Forest (RF) is an increasingly popular machine learning method.[34] It builds an ensemble of many decision trees trained separately on a bootstrapped sample set of the original data. Each decision tree grows by randomly selecting a subset of candidate attributes for splitting at each node. For our application we use 200 trees for training our classifier on each set of image features.

Support vector machines (SVM) is one of the most widely used classifiers in medical image analysis.[31, 35] It performs classification by constructing a hyperplane in a high-dimensional feature space. It can use either linear
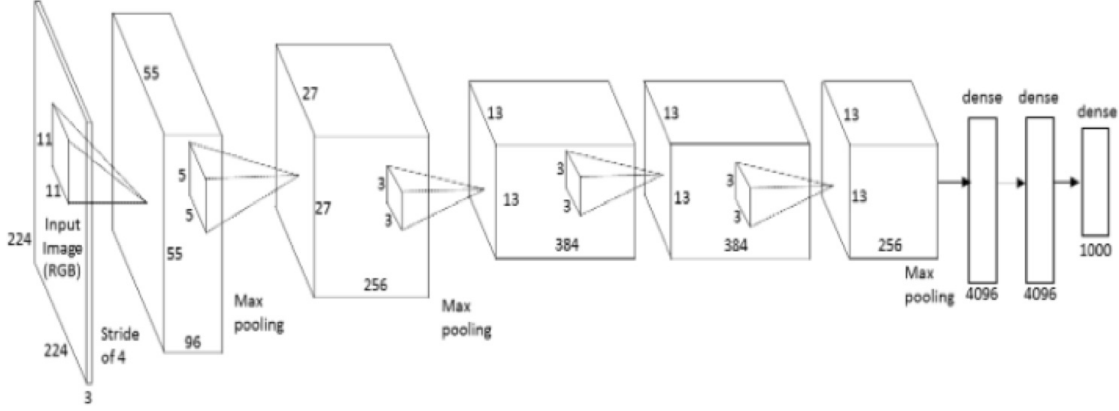
Figure 2. Visualization of a Deep Convolutional Neural Network as described in Krizhevsky et al.[11] We modify and utilize the above network to extract features and also perform classification. The modifications still maintain the benefits of the pretrained network for our cytopathology application.

or nonlinear kernels, and its effectiveness depends on the selection of kernel, and the kernel's parameters. Linear SVM is widely used because it has good performance and is relatively fast in its classification task. We utilize Linear SVMs and standardize the data by the weighted column mean and standard deviation. Our SVM is a multiclass model with binary SVM learners.

k-Nearest Neighbors (kNN) is one of the simplest classifiers, which classifies a new instance by a majority vote of its k nearest neighbors. In this paper, we use the $\chi^2(i,j) = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k)-h_j(k)]^2}{h_i(k)+h_j(k)}$, distance metric to find the k nearest neighbors. Here, K is equal to the number of bins for the different image features, h. We use $k = 5$ in our experiments.

Convolution Neural Networks (CNN) - Building a convolutional neural network is a data intensive and time consuming task. For example, AlexNet,[11] shown in Figure 2 takes several weeks to train on multiple GPUs on the ImageNet[36] database. Since a deep neural network can contain millions of weight parameters, it is essential that you have enough data to train these parameters without overfitting. In our scenario, we do not have millions of labeled medical images and so we take a different approach called "fine-tuning".[37] We replace the last two layers of our with newly initialized weights, and change the output of the last layer to 3 (three different stain types). The other layer weights in the network remain fixed from previous training on ImageNet. The last layer remains a soft-max classification probability, $\hat{p}_n$, as defined by the following equation,

$$\hat{p}_k = \frac{e^{w_k}}{\sum_{k'=1}^{K} e^{w_{k'}}} \tag{1}$$

where k is a class in K total classes and w represents the output of the last layer of our network, $w \in [-\infty, +\infty]$.

We fine tune the network with the the following parameters: base learning rate = 0.001, gamma = 0.1, over 2000 iterations. The momentum is 0.9 with a weight decay of 0.0005.

4. RESULTS

4.1 Mobile Application

For development of our mobile front end, we use the popular cross-platform tool, PhoneGap. The mobile application consists of several different sections that allow you to learn about different types of thyroid disease, and browse the thyroid ontology. The content is dynamically generated through various AJAX and SPARQL calls to a backend database of webpages and RDF triple content. We are able to navigate through the ontology and display information from our custom knowledge representation or we can display data linked through the Semantic Web capabilities of our framework, see Figure 3.

(a) Screenshot 1   (b) Screenshot 2   (c) Screenshot 3   (d) Screenshot 4   (e) Screenshot 5   (f) Partial Ontology
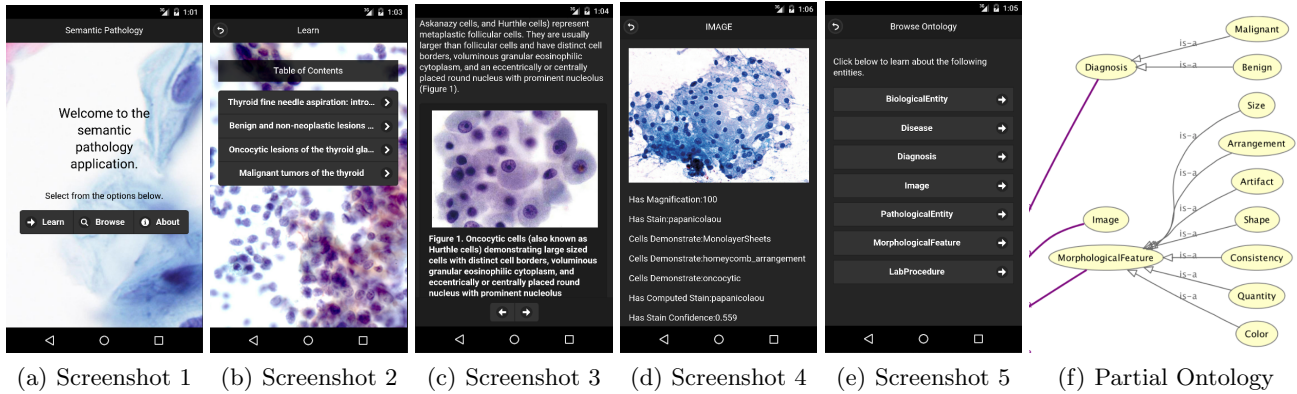
Figure 3. The mobile screen shots (a-e) illustrate our application capabilities. Visualization of part of our ontology (f) and corresponding screen shots of our application navigating the ontology (d-e).

4.2 Image Property Classification Results

Give the image features and machine learning methods, we can evaluate their performance, pick the best combination and enhance our ontology with the automatically computed image metadata. We experimented on a single classification experiment, automatically recognizing the stain type of an image, and record the computed class and confidence using the given image properties.

hasComputedStain - We train four classifiers on four image feature types and present our accuracy results in Table 1. The four machine learning algorithms used were Random Forest (RF), Linear Support Vector Machines (SVM), K-nearest neighbors (KNN), and Convolutional Neural Networks (CNN). The four image features that we utilized for the experiments were PLAB (color features), PHOG and PLBP (texture features), a combination of color and texture features where PLAB, PHOG, and PLBP are combined into one feature vector, and finally the CNN codes. We experiment on a dataset of 459 labeled cytology images, which consists of 391 training images and 68 testing images. The data consists of a closely balanced set of three stain classes, H&E, Romanowsky, and Papanicolaou. Although the classes are balanced, the pathological variation within the images is highly varied, including different disease, cell types, levels of abnormality, and magnification. The variability of the dataset can be seen in Figure 4.

hasComputedStainConfidence - Based upon our final accuracies, (see Table 1) we used the CNN codes in a linear SVM machine learning algorithm for the stain confidence rating. The confidence of the method is reported in the image property as a posterior probability of the computed stain.

In the limited dataset that we are working with, it is presumptuous to make any definitive conclusions; however, we can make some interesting observations. First of all, it is clear that the use of CNN codes over any hand crafted feature makes a significant difference in classification rates. Secondly, it appears that the machine learning algorithm is not all that important. In fact, RF, SVM, and CNN machine learning methods all do about the same accuracy on the test set (approx. 70%) as long as the features used for classifications are the ones generated through the deep neural network.

4.3 Future Work

Our work in building a thyroid pathology ontology is ongoing. We are continually enhancing and revising the ontology towards a more comprehensive system. Our immediate technological plan is to improve the user interface so that it is more friendly for the non-expert. We also plan to ensure the robustness of the application in regard to scalability. A future goal for the application is to build a diagnostic component to the application where a pathologist would be able to traverse from the initial visual inspection of an unknown FNA sample, upload the image to our platform, run the image through our machine learning framework, and follow a vetted flowchart to come to a similar image or patient to assist with the diagnosis. We plan to continue work in the area of machine learning and extend our framework to look at more disease specific characteristics in cytopathology, segmentation of cell structures, as well as explore mutli-label classification tasks.
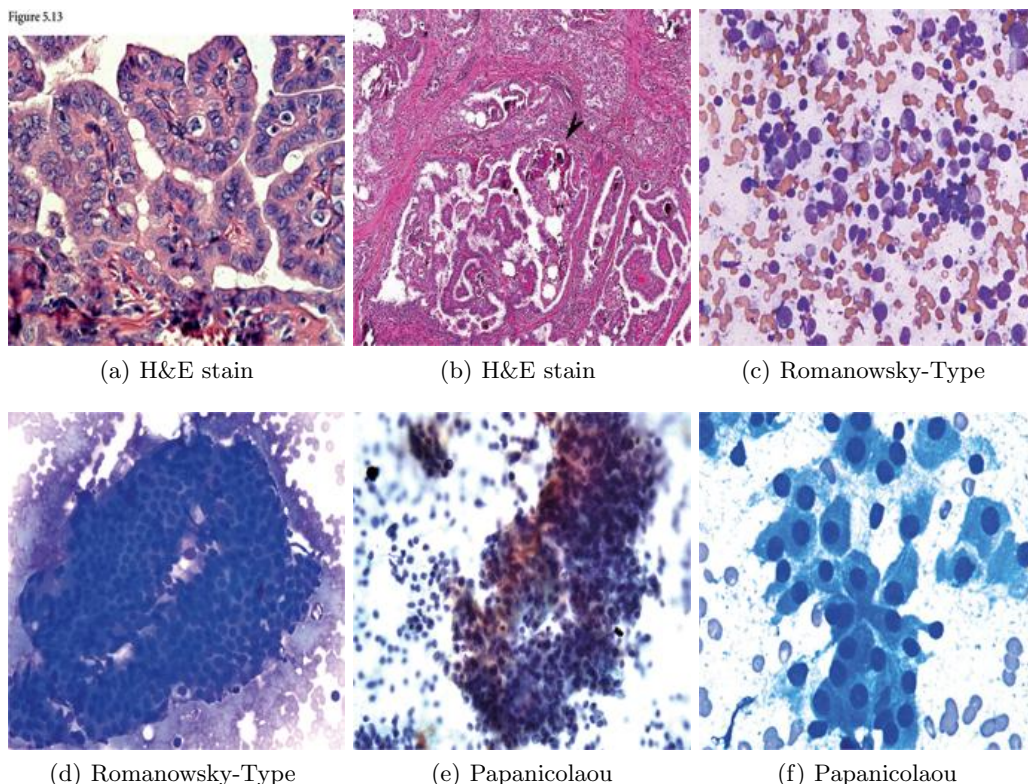
Figure 4. Sample images from our training and test set. From these images, one can see the variability of color and texture contained in the different types of stain preparations. Images obtained from Baloch et al.[3]

REFERENCES

[1] Hughes, D. T., Haymart, M. R., Miller, B. S., Gauger, P. G., and Doherty, G. M., "The most commonly occurring papillary thyroid cancer in the united states is now a microcarcinoma in a patient older than 45 years," *Thyroid* **21**(3), 231–236 (2011).

[2] Cooper, D. S., Doherty, G. M., Haugen, B. R., Kloos, R. T., Lee, S. L., Mandel, S. J., Mazzaferri, E. L., McIver, B., Pacini, F., Schlumberger, M., et al., "Revised american thyroid association management guidelines for patients with thyroid nodules and differentiated thyroid cancer: the american thyroid association (ata) guidelines taskforce on thyroid nodules and differentiated thyroid cancer," *Thyroid* **19**(11), 1167–1214 (2009).

[3] Baloch, Z., Elsheikh, T., Faquin, W., and Vielh, P., [*Head and Neck Cystohistology*], Cambridge University Press (2014).

[4] Ali, Y. Z., Nayar, R., Krane, J. F., and Westra, W. H., [*Atlas of Thyroid Cytopathology: With Histopathologic Correlations*], Demos Medical Publishing (2013).

[5] Kini, S. R., [*Thyroid cytopathology: an atlas and text*], Lippincott Williams & Wilkins (2008).

[6] "Unified medical language system [accessed august 2015]," (2015). `http://www.nlm.nih.gov/research/umls/`.

[7] Sioutos, N., de Coronado, S., Haber, M. W., Hartel, F. W., Shaiu, W.-L., and Wright, L. W., "Nci thesaurus: a semantic model integrating cancer-related clinical and molecular information," *Journal of biomedical informatics* **40**(1), 30–43 (2007).

[8] Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., and Hellmann, S., "Dbpedia-a crystallization point for the web of data," *Web Semantics: Science, Services and Agents on the World Wide Web* (2009).

Stain	Feature	5KNN % acc	RF % acc.	Linear-SVM % acc.	Deep CNN % acc.
Papanicolaou	Color	65.3	**92.3**	65.3	n/a
Papanicolaou	Texture	7.6	11.5	**34.6**	n/a
Papanicolaou	Color+Texture	**61.5**	30.7	**61.5**	n/a
Papanicolaou	CNN code	57.6	76.9	**84.6**	65.3
Romanowsky	Color	19.0	23.8	**42.8**	n/a
Romanowsky	Texture	28.5	33.3	**47.6**	n/a
Romanowsky	Color+Texture	38.1	23.8	**47.6**	n/a
Romanowsky	CNN code	**52.3**	38.1	**52.3**	47.6
H&E	Color	**85.7**	47.6	61.9	n/a
H&E	Texture	**95.2**	**95.2**	**95.2**	n/a
H&E	Color+Texture	90.4	**95.2**	**95.2**	n/a
H&E	CNN code	61.9	90.4	71.4	**95.2**
Avg All Stain	Color	57.3	**57.4**	**57.4**	n/a
Avg All Stain	Texture	41.1	44.1	**57.3**	n/a
Avg All Stain	Color+Texture	63.2	48.4	**67.6**	n/a
Avg All Stain	CNN code	57.2	69.1	**70.5**	69.1

Table 1. Accuracy of four machine learning algorithms on four extracted image features. The last rows display the average of the method over all three stain classes.

[9] "Papanicolaou society of cytopathology [accessed august 2015]," (2015). http://www.papsociety.org/atlas.html.

[10] "Nci bethesda system [accessed august 2015]," (2015). http://nih.techriver.net/.

[11] Krizhevsky, A., Sutskever, I., and Hinton, G. E., "Imagenet classification with deep convolutional neural networks," in [*Advances in neural information processing systems*], 1097–1105 (2012).

[12] Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., and Darrell, T., "Decaf: A deep convolutional activation feature for generic visual recognition," *arXiv preprint arXiv:1310.1531* (2013).

[13] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., and Darrell, T., "Caffe: Convolutional architecture for fast feature embedding," in [*Proceedings of the ACM International Conference on Multimedia*], 675–678, ACM (2014).

[14] LeCun, Y., Kavukcuoglu, K., and Farabet, C., "Convolutional networks and applications in vision," in [*International Symposium on Circuits and Systems (ISCAS)*], 253–256 (2010).

[15] Horvat, M., Gledec, G., and Sjekavica, T., "Advantages of semantic web technologies usage in the multimedia annotation and retrieval," *International journal of computers and communications* **8**, 41–48 (2014).

[16] Kim, E., Huang, X., and Heflin, J., "Finding vips-a visual image persons search using a content property reasoner and web ontology," in [*IEEE International Conference on Multimedia and Expo (ICME)*], 1–7 (2011).

[17] Van Woensel, W., Al Haider, N., Roy, P. C., Ahmad, A. M., and Abidi, S. S., "A comparison of mobile rule engines for reasoning on semantic web based health data," in [*IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*], **1**, 126–133, IEEE (2014).

[18] Tolksdorf, R. and Bontas, E. P., "Engineering a semantic web for pathology," in [*International Wireless Communications Expo*], **2004**, 585–586 (2004).

[19] Sertel, O., Lozanski, G., Shana'ah, A., and Gurcan, M. N., "Computer-aided detection of centroblasts for follicular lymphoma grading using adaptive likelihood-based cell segmentation," *IEEE Transactions on Biomedical Engineering* **57**(10), 2613–2616 (2010).

[20] Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., and Yener, B., "Histopathological image analysis: A review," *Biomedical Engineering, IEEE Reviews in* **2**, 147–171 (2009).

[21] "Phonegap [accessed august 2015]," (2015). http://phonegap.com.

[22] "Apache jena [accessed august 2015]," (2015). http://jena.apache.org.

[23] "Jquery mobile [accessed august 2015]," (2015). https://jquerymobile.com/.

[24] Dalal, N. and Triggs, B., "Histograms of oriented gradients for human detection," in [*IEEE Computer Society Conference on Computer Vision and Pattern Recognition*], **1**, 886–893 (2005).

[25] Rodenacker, K. and Bengtsson, E., "A feature set for cytometry on digitized microscopic images," *Analytical Cellular Pathology* **25**(1), 1–36 (2003).

[26] Palcic, B., MacAulay, C. E., Harrison, S. A., Lam, S., Payne, P. W., Garner, D. M., and Doudkine, A., "System and method for automatically detecting malignant cells and cells having malignancy-associated changes," (2000).

[27] Kuncheva, L. I. and Whitaker, C. J., "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Machine learning* **51**(2), 181–207 (2003).

[28] Doyle, S., Hwang, M., Shah, K., Madabhushi, A., Feldman, M., and Tomaszeweski, J., "Automated grading of prostate cancer using architectural and textural image features," in [*IEEE International Symposium on Biomedical Imaging*], 1284–1287 (2007).

[29] Doyle, S., Rodriguez, C., Madabhushi, A., Tomaszewski, J., and Feldman, M., "Detecting prostatic adeno-carcinoma from digitized histology using a multi-scale hierarchical classification approach," in [*International Conference of the IEEE Engineering in Medicine and Biology Society*], 4759–4762 (2006).

[30] Deng, L., Li, J., Huang, J.-T., Yao, K., Yu, D., Seide, F., Seltzer, M., Zweig, G., He, X., Williams, J., et al., "Recent advances in deep learning for speech research at microsoft," in [*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*], 8604–8608 (2013).

[31] Kim, E. and Huang, X., "A data driven approach to cervigram image analysis and classification," *Color Medical Image analysis* **6**, 1–13 (2013).

[32] Bosch, A., Zisserman, A., and Munoz, X., "Representing shape with a spatial pyramid kernel," in [*CIVR*], 401–408 (2007).

[33] Vedaldi, A. and Fulkerson, B., "VLFeat: An open and portable library of computer vision algorithms." http://www.vlfeat.org/ (2008).

[34] Friedman, J., Hastie, T., and Tibshirani, R., [*The elements of statistical learning*], vol. 1, Springer series in statistics Springer, Berlin (2001).

[35] Morra, J. H., Tu, Z., Apostolova, L. G., Green, A. E., Toga, A. W., and Thompson, P. M., "Comparison of adaboost and support vector machines for detecting alzheimer's disease through automated hippocampal segmentation," *IEEE Transactions on Medical Imaging* **29**(1), 30–43 (2010).

[36] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L., "Imagenet: A large-scale hierarchical image database," in [*Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*], 248–255, IEEE (2009).

[37] Yosinski, J., Clune, J., Bengio, Y., and Lipson, H., "How transferable are features in deep neural networks?," in [*Advances in Neural Information Processing Systems*], 3320–3328 (2014).