Object Recognition and Active Learning in Microscope Images

Conor Nugent and Pádraig Cunningham

Department of Computer Science
Trinity College Dublin
Conor.Nugent@cs.tcd.ie
Padraig.Cunningham@cs.tcd.ie

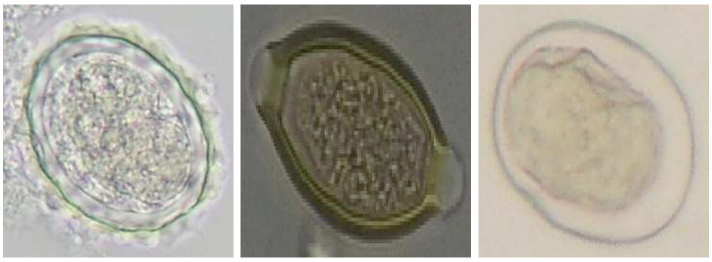
Abstract. Microscopic analysis forms an integral part of many scientific studies. It is a task which requires great expertise and care. However, it can often be an extremely repetitive and laborious task. In some cases many hundreds of slides may need to analysed, a process that will require each slide to be meticulously examined. Machine vision tools could be used to help assist in just such repetitive and tedious tasks. However, many machine vision solutions involve a lengthy data acquisition phase and in many cases result in systems that are highly specialised and not easily adaptable. In this paper, we describe a framework that applies flexible machine vision techniques to microscope analysis and utilises active learning to help overcome the data acquisition and adaptability problems. In particular we investigate the potential of various aspects of our proposed framework on a particular real world microscopic task, the recognition of parasite eggs.

1 Introduction

The microscope is an irreplaceable tool for analysis in many disciplines of science. It is used in a wide variety of vastly differing tasks each of which require specific expertise. Although this work requires skill and expertise it can often be repetitive and laborious work. Machine-vision techniques could be used to help assist in some of these tasks and this has been done for some specific tasks [1] [2] [3]. However, such solutions can suffer from the same lengthy and difficult data acquisition and model building phases that hamper machine vision solutions generally. The solutions offered in such cases are often highly specific and not easily adaptable or updated. To overcome these problems we propose a novel framework which utilises an adaptable recognition technique along with active learning to reduce the data acquisition process and leverage expert knowledge in a direct but efficient manner. The framework should serve as an interactive assistant to its user and continually update itself based on its performance and on new information received.

The particular problem we have addressed is the classification of parasite eggs. We had previously considered histopathology images, however, the large variation in appearance that many similar conditions could have, made this

Fig. 1. An example of section of a slide



a. Ascuris b. Trichuris c. Ringworm

Fig. 2. Examples of parasite eggs

an unfeasible machine vision task. Parasite eggs, however, have a well defined structure which makes them amenable to recognition tasks. We collected a large set of images from a set of samples which had been taken from human subjects. In Figure 1 a small section of one image of a microscope slide can be seen. Typically a parasitologist may have to sift through many hundreds of slides searching for such objects as those seen in Figure 2. This can be an extremely tedious task but one which none the less must be carried out methodically

We chose this task as a testing ground for many aspects of our proposed framework as it highlights the many challenges facing the application of machine-vision and active learning in such domains as well as the potential rewards. Clearly if a recognition system could be devised it would be very useful.

However, the data acquisition task is a formidable one with many different parasite eggs to identify and each having many variations in appearance. There is also a large presence of irrelevant objects in slide images which are of no interest and must be ignored. This makes the automation of this process through active learning desirable as it can be used to harness expert opinion in an efficient way given the large volumes of unlabelled, irrelevant and nosy data. Since mistaken classifications in such a task would be highly damaging the probabilistic and adaptive approach taken by active learning means that speculative classifications could be referred for expert opinion and the system updated. Ideally we would also like our system to be easily updated with new classes of objects such as new parasite egg varieties in this example.

In the next section we will outline the structure of our proposed framework before describing particular aspects we have implemented in greater detail. In Sections 2.1 and 2.2 we describe the recognition and active learning algorithms we have used. We then evaluate performance of these key aspects of the framework in Section 3. Finally we end with some conclusions from our work.

2 Framework

The framework we have developed has a number of distinct phases to its operation. We will now explain each of these in turn:

1. **Image Acquisition:** Firstly the microscopic images are captured in digital format.
2. **Preprocessing:** The microscope images as seen in Figure 3 contain large empty regions. Ideally before we start processing these images we would like to eliminate large uninteresting areas and pick the objects of interest. One simple way in which we can eliminate the background and pick out individual objects is by using a series of standard morphological image processing techniques. This process is displayed in Figure 3. The operations are carried out on the image in Figure 3.a and the results can be seen in Figure 3.b. The background has been eliminated and collections of pixels proposed to be related are colour-matched. By eliminating groups of pixels too small to be of interest we then crop the remaining objects from the image as depicted in Figure 3. This leaves use with a set of templates to process.
3. **Image Recognition and Identification:** Clearly a large amount of unlabelled data will still remain which we would like to build our recognition system on. We have access to a domain expert whom we can query examples with. However, we would like to reduce the workload on the domain expert as much as possible and build an accurate recognition system quickly. To do this we use active learning which we outline in greater detail in Section 2.2.

a. before processing b. objects identified c. salient objects cropped

Fig. 3. Examples of parasite eggs

We have briefly outlined our framework. We will now go into greater detail on two key facets; the recognition and active learning algorithms.

2.1 Image Recognition: EigenImages

In choosing an image recognition technique there were a number of requirements stemming from the role of the recognition system within our framework that needed to be fulfilled. These requirements can be summarised as follows:

- **Generalise:** the framework is designed to be adaptable and applicable to a broad range of problems. Many machine vision solutions make assumptions about specific characteristics of the objects to be recognised which are incorporated into the recognition algorithm. We wished to avoid including problem specific assumptions as much as possible as these limited the flexibility of our approach and weren't in fitting with the spirit of active and continuous learning.
- **Powerful and Quick:** Any recognition system to be used in a active learning context needs to be reasonably quick and easily updated based on new information received.
- **Probabilistic:** Given a large set of unlabelled data many active learning algorithms use probabilistic measures of the likely class on an unknown object to help inform which objects to consult a domain expert about. This characteristic can also be used to ensure that doubtful classification are not made.
- **Perform Identification and Recognition:** Before recognition we first need to locate the objects of interest which may be scattered across large images. As we saw in Section 2 the pre-processing stage involves detecting salient objects but doesn't differentiate between parasites and other objects. If the recognition technique we are using can be used in pick out the objects of interest in a more direct way this would help remove many irrelevant objects at an early stage.

We considered a broad range of possible recognition techniques including adaptive-subspace self-organising maps (ASSOMs) and techniques used in image retrieval system such as RETIN [4] [5] [6]. However, we found Eigenimage-based recognition techniques best meet the criteria we had set for our recognition system. In the next Section we will very briefly describe the basis of this approach.

Eigenimages The Eigenimages approach originally grew out of research by Turk and Pentland into automated face recognition and detection in images [7]. Much of the previous work in face recognition up until that time had been based on assumptions about what features are important to facial recognition such as eyes, noses and lips. These assumptions meant that such approaches were sometimes fragile and difficult to adapt to problems which were only subtly different. It had also been shown that such features and their relationships didn't account for every aspect of the performance of adult human recognition [8].

Turk and Pentland applied an information theory approach to extracting out what the salient and the information features were. They wished to encode a set of facial images as efficiently as possible independent of any features. In doing this they were able to determine a set of features that accurately defined the variations and differences of the faces in their image set. The process used to create such an encoding is Principle Component Analysis (PCA) [9]. A covariance matrix of the set of images is calculated treating each image as a vector in a high dimensional space (defined by the number of pixels). The eigenvectors and corresponding eigenvalues are then calculated from this matrix. These eigenvectors can then be thought of as a set of features each of which describes a variation which exits in the image data. In practice only a small set of the eigenvectors are informative and eigenvectors with small eigenvalues can be ignored. Each image can be described as a linear combination of the eigenvectors and each image is projected into this greatly reduced dimension space. In the case of facial recognition this space is referred to as the *face-space* and each eigenvector as an eigenface. In our case the space is a *parasite-space* and each eigenvector as a *eigen-parasite*. Each known image is projected into parasite-space and is defined as a linear combination of eigen-parasites. Classification can be achieved by projecting the unknown image into the parasite-space and finding the known image that matches its projection most closely. Eigenimage-based recognition continues to be an active area of research with many extensions involving different learning models to this basic approach being developed. [10]. Bearing in mind the requirements we set out in Section 2.1 we decided to use a weighted nearest neighbour algorithm.

2.2 Active Learning

Active learning is a term used to describe a learning method where the classifier has control over the training data it uses [11]. This is different to the traditional supervised learning approach where the classifier has no control over its learning process and is given a fixed set of labelled training data. An active learner makes use of an external teacher or domain expert who will label a presented example if the active learner requests them to. Given a large set of unlabelled examples the active learner tries to select and present to the expert the most informative examples which will maximise its learning and performance. Active Learning is particularly useful in domains were there is an abundant amount of raw unlabelled data available but much of data is irrelevant, uninformative and the costs of labelling that data are high. In the next Section we will outline the particular active learning algorithm we used in our framework, uncertainty-based selection.

Uncertainty-based Approach Uncertainty-based sampling is a active learning method for query selection which utilises a single classifier to make an informed selection from a set of unlabelled examples [12]. The method uses the uncertainty that the classifier has about the unlabelled examples in the pool in order to choose queries. The idea of uncertainty stems from the confidence the classifier has in predicting a label for a given unlabelled example. For this reason a probabilistic classifier is needed. The example that the classifier is least uncertain about is deemed to be the one it would gain greatest benefit from being labelled and so is selected. The algorithm is quite simple and is outlined in the figure below.

Given:
- U: a large set of unlabelled data
- S: a small set of initially labelled data
- L: a probabilistic leaner
- E: a domain expert
- SE: selected example

Algorithm:
- 1- build the learner, L, on S
- 2- use L to compute the probability of each example in U
- 3- select the example with the lowest probability as SE
- 4- label SE using E and add to S
- 4- remove SE from U
- 6- repeat until stopping condition is met

The issue of what stopping condition to use is an open question in the active learning community and most solutions are ad hoc and domain dependent. Another issue is the selection of examples to seed the algorithm. In the case of evaluations we carry out in the section we seed the active learner with three examples chosen at random from the data set.

3 Evaluations

In order to evaluate our framework in this stage of its development we chose to evaluate two keys facets, the image recognition accuracy and the effectiveness of active learning in image data. We created a dataset of images of three different parasite eggs; *trichuris*, *ascuris* and *ringworm*. The dataset contained 59 images in total, distributed between each of the classes as shown in Table 1. The images were cropped similarly to those shown in Figure 2 however many were out of focus, badly aligned or occluded. In the next two sections we will outline the experimental method and results of our evaluations using this data

3.1 Recognition

In order to evaluate the recognition potential of our recognition system we performed 10 fold cross-validation. Initially, using the raw images, the results ap-

Table 1. Summary of distribution of classes in the parasite data

Parasite Type	No. of Examples
trichuris	15
ascuris	22
ringworm	22

peared to be extremely poor at just 62%. This was a long way short of levels of accuracy we have achieved on facial recognition tasks. Typically on such tasks we achieved accuracies greater than 95% (as can be seen in Section 3.2). However, we had overlooked two important characteristics that differ between the parasite data and face data.

Firstly, face data is typically well aligned. While lighting conditions and subject angles may differ the images usually do not deviate from a typical mug-shot set-up. Images of people faces upside down or at a horizontal angle are not encountered. This is not the case in the parasite data. The parasites are not restricted in this same way and can be seen at many different orientations. This makes the recognition task far harder as the same parasite egg will look quite different to a recognition system if it is encountered at completely different orientations. To combat this problem we simply rotated each image in the training set to different angles and included these variations in the training set too. This led to a drastic improvement in accuracy. As be seen in Table 2 we were able to boost the accuracy up 15% to 77.84%.

Table 2. Summary of Accuracy results for different pre-processing techniques

Data	Percentage Accuracy
unaltered	62.29
masked	73.75
rotated	77.84
masked and rotated	81.96

The second problem was slightly more subtle. As Turk and Pentland observe "the background can significantly effect the recognition performance, since the eigenface analysis ..does not distinguish the face from the rest of the image". In order to combat this we *masked* our images by applying a simple morphological image preprocessing technique to help eliminate background detail. This also greatly improved the accuracy achieved increasing it to 73.75%. Using these two processing techniques together ultimately led to approximately a 20% increase in accuracy to 81.96%.

Although 81.96% represents a reasonable performance it is still short of that which would be needed for a real world working system. However, as we stated before, we included many out of focus and occluded examples which were bound to impact on accuracy performance. The problems posed by the inclusion of such examples can be overcome and our accuracy figure represents a good baseline that can easily be improved upon. As we shall see in the next section one such way is the intelligent selection of training data.

3.2 Active Learning

We implemented the uncertainty-based selection algorithm as outlined in Section 2.2. In order to evaluate this scheme we needed to create a training set of data to select images from and also a separate test set to evaluate the effect that each image selection has on accuracy. We divided our data into 5 sections and carried out an evaluation process similar to that of cross-validation. In each iteration 4 sections were used as the training set and the fifth as a test set. This is repeated 5 times until each section of data has served as a test set. As a baseline with which to compare uncertainty-based selection we performed the same evaluation with a selection strategy in which images are selected at random. The results for the two selection schemes over the 5 evaluations were then averaged and can be seen in Figure 4.

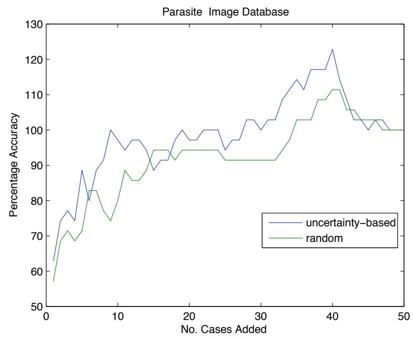


Fig. 4. Uncertainty-based v's. random selection for parasite data

The graph displays the accuracy after each case is labelled as a percentage of the accuracy that could be achieved using all the data. Uncertainty-based selection clearly outperforms random selection. In fact uncertainty-based selection manages to achieve the same level of accuracy as that achieved on a full image set with less then 10 images needing to be labelled. Interestingly, uncertainty-based and random selection both achieved levels of accuracy higher than the baseline with uncertainty-based selection achieving accuracy rates that are 120% improvements on using the entire image set. This reflects the fact that the data contains images that are poor and highlights the strengths of active learning in eliminating such images.

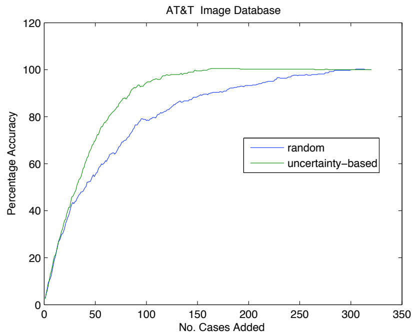


Fig. 5. Uncertainty-based v's. random selection for AT&T data

We also wished to evaluate whether such a performance was typical on image data or characteristic of the small image-set we had. We tested uncertainty-based selection on a face recognition task. In particular we used the AT&T face image set which has been kindly made available by AT&T Laboratories Cambridge. This image set contains 400 images of 40 individuals. Our 10 fold cross-validation accuracy for this image set is 97.25%. We performed the same evaluation as before and the results can be seen in Figure 5. Again uncertainty-based selection out performs random selection and baseline accuracy is achieved with less than 150 images being labelled which is just 40% of the image-set. Here too uncertainty based selection leads to improvements over the base accuracy set by using all the image data. However, the improvements are very modest reflecting the quality of the image data and the small margin for improvement.

3.3 Conclusions from evaluations

It is clear from our evaluations that uncertainty-based sampling can achieve good and sometimes improved levels of accuracy with far fewer cases needing to be labelled. This justifies the use of active learner within our framework. An accurate classifier can be built with a greatly reduced work load being placed upon the domain expert. However in a real world application the question of when to stop the active learning process will have to be addressed. This is an outstanding question in the active learning community however it is possible that a domain specific solution could be devised.

We have also demonstrated that our chosen recognition system is accurate and works well within an active learner framework. The recognition system is also adaptable, we have used the same recognition system for face data and parasite images. Although, small adjustments needed to be made to the recognition system to apply it to the parasite data these adjustments were not specific to just the parasite data but apply to all such objects viewed through a microscope.

Eigenimage recognition methods have been used to detect objects such as faces and hands in images. This is still something we have to evaluate on microscope image data. The addition of this functionality would greatly improve our segmentation process and the quality of our images.

4 Conclusion

The application of machine vision techniques to microscopic image data is a potentially rewarding but difficult task. We have proposed a framework for image recognition that utilises active learning to help overcome these difficulties. Key aspects of our design have been implemented and evaluated in the challenging parasite recognition domain. Clearly there is still much work needed and many problems to be addressed if a working application is be realised. We have, however, demonstrated in principle that our framework has potential and in particular that active learning can help alleviate the image acquisition problem in machine vision.

References

1. Carrión, P., Cernadas, E., Gálvez, J.F., Damián, M., de Sá-Otero, P.: Classification of honeybee pollen using a multiscale texture filtering scheme. Mach. Vision Appl. **15** (2004) 186–193
2. Perner, P., Günther, T.: Detection of hygiene-relevant parameters from cereal grains based on intelligent image interpretation and data mining. In Bauer, M., Brandherm, B., Fürnkranz, J., Grieser, G., Hotho, A., Jedlitschka, A., Kröner, A., eds.: LWA, DFKI (2005) 216–219
3. Perner, P., Perner, H., Jänichen, S., Bühring, A.: Recognition of airborne fungi spores in digital microscopic images. In: ICPR (3). (2004) 566–569
4. Fournier, J., Cord, M., Philipp-Foliguet, S.: Retin: A content-based image indexing and retrieval system. Pattern Analysis and Applications **4** (2001) 153 – 173
5. Zheng, H., Liu, H., Daoudi, M.: Blocking objectionable images: adult images and harmful symbols. In: ICME, IEEE (2004) 1223–1226
6. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vision **40** (2000) 99–121
7. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience **3** (1991) 71–86
8. Carey, S., Diamond, R.: From piecemeal to configurational representation of faces. Science **195** (1977) 312–314
9. Jolliffe, I.T.: Principal Component Analysis. Springer (1986)
10. Moghaddam, B., Pentland, A.: Probabilistic visual learning for object representation. IEEE Transactions on Pattern Analysis and Machine Intelligence **19** (1997) 696–710
11. Davy, M.: A review of active learning and co-training in text classification. Technical report, Computer Science Dept., Dublin University, Trinity College (2005)
12. Lewis, D.D., Gale, W.A.: A sequential algorithm for training text classifiers. In: SIGIR '94: Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval, New York, NY, USA, Springer-Verlag New York, Inc. (1994) 3–12