Using Transfer Learning for High Throughput Toxin Screening

Patrick Hayes, Aaron Graham

Abstract

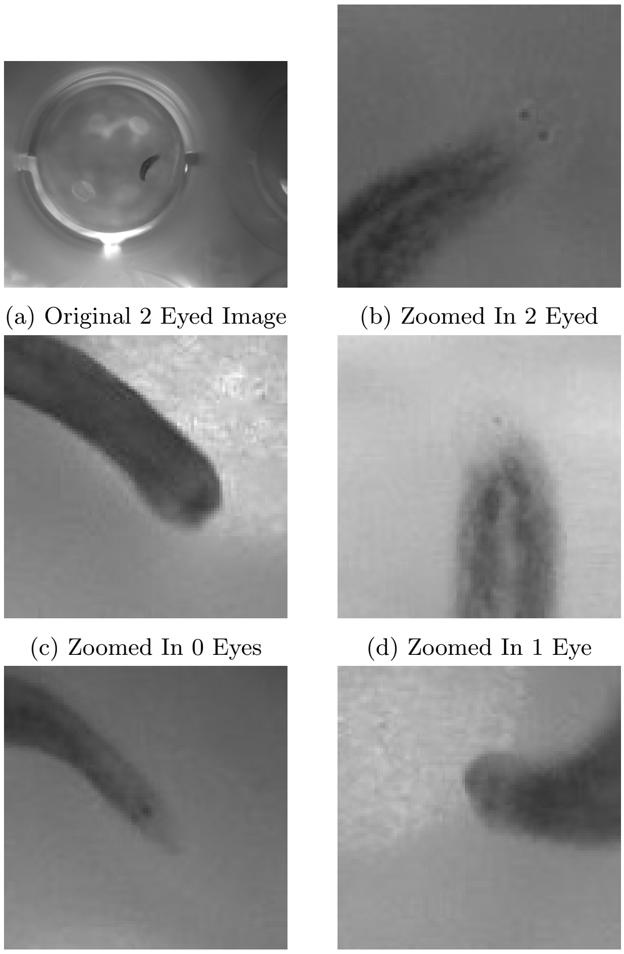
Toxicology – the branch of science concerned with the nature, effects, and detection of poisons – has traditionally relied on expensive mammalian studies[1]. However, due to the large number of environmental toxins that need testing, less expensive, high throughput alternatives are required. Planaria – a small asexual flatworm – provide an inexpensive and scalable solution[1]. The small size of planaria not only makes them inexpensive to maintain, but also easy to image with a high-resolution camera. Researchers can run many test simultaneously, record the planarias reaction, and use computer vision techniques to analyze the results in a cost-effect and timely manner. The results of some of these tests can be analyzed using traditional computer vision techniques; however, many tests involve classifying features of the planaria – such as their body shape or eyes – as normal or abnormal. Deep convolutional neural networks have achieved state of the art results on image classification tasks, but they require a large training set of images[2]. Researchers do not have the time or resources to hand-label sufficient images for this technique to be effective. Using transfer-learning, we have developed a software platform for researchers to quickly classify images with near state-of-the-art performance and minimal training data.

1 Classifying Eye Development

This paper will demonstrate the effectiveness of using transfer learning to identify whether a toxin has impaired a planarian's ability to develop eyes. The following paragraphs provide information about the images being classified, describe how the images should be classified, and list metrics of performance.

1.1 Data

The images that are being classified come in short videos of 50 frames each. Each video observes a single planarian swimming in a well for 5 seconds. Each frame is a gray-scale image with 1280 x 1024 pixels. Because of planaria's small size even with the use of a high resolution camera, their eyes are only a small cluster of pixels ranging from around 3-10 square pixels. The eyes of planaria can also vary in shape and color relative to their body. Sometimes, the planarian



(a) Original 2 Eyed Image (b) Zoomed In 2 Eyed (c) Zoomed In 0 Eyes (d) Zoomed In 1 Eye (e) Profile View (f) Slanted View

Figure 1

is on it side or at an angle when the photos are taken which makes the eyes look different. When a toxin impairs a planarian's ability to develop eyes, the planarian can either grow no eyes at all or only grow a single eye, Figure 1 shows example images. There are thousands of unlabeled video clips, but less than 5% of the video are of planaria that have been impaired by toxins.

1.2 Standards For Classification

Each short video clip will be classified as either abnormal or normal where abnormal and normal are defined as follows. An abnormal video observes a planarian with a single eye or no eyes at all and a normal video observes a planarian with two eyes. The eyes of the planarian may be hidden from view for much of the video – due to the orientation of the worm or poor lighting – but if both eyes are clearly visible in a single frame the video is normal. If the video provides insufficient information to classify a planarian, the video should be labeled unknown.

1.3 Metrics of Performance

1.3.1 False Negative Rate

The consequences of falsely labeling a toxic chemical as safe are much more severe than falsely labeling a harmless chemical as toxic. If a harmless chemical is falsely labeled toxic it will go on to further testing where its true nature will be discovered; however if a harmful chemical is falsely labeled safe, it will likely go to market where its effects could be catastrophic. As such, reducing the false negative rate has a much higher priority than reducing the false positive rate, although an effective solution will have a low error rate for both.

1.3.2 Recall

To ensure low error rates, an effective solution may choose to classify certain planaria as unknown. If a potential solution labels less than 10% of videos as unknown it is still a viable solution, but solutions with higher recall are favored.

1.3.3 Time and Computational Resources Needed

Faster solutions are favored, but at the lowest priority. As long as a solution requires minimal human supervision and achieves low error rates, it is considered effective even if it takes an hour to process 50 images.

2 Why Use Transfer Learning?

Deep convolutional neural networks have achieved state of the art performance on image classification tasks but require training sets containing millions of images[2]. Transfer learning allows researchers with limited training data to make use of the feature extractors from a pre-trained deep network[3]. Many of these feature extractors generalize across classification space; a system that classifies between cats and dogs and a system that classifies between apples and bananas, both require edge detectors. After using the feature extractors of a deep well network, researchers need only train a small densely connected network to classify their images based off these pre-computed set of features. For more complex classification tasks, researchers may also need to train their own set of high-level feature extractors to extract abstract features specific to their classification task, but for the classifying eye development task we were able to achieve low error rates with out additional convolutional layers.

3 Why Use Neural Networks?

While neural networks have achieved state of the art accuracies in image classification challenges, they do have some drawbacks. Even using transfer learning, neural networks still require hand labeling thousands of images. Furthermore neural networks are computational expensive compared to other non-statistical computer vision techniques. However, a neural networks ability to handle large variations in the input images justifies these costs.

The eyes of a planarian tend to be much darker than the surrounding tissue of its head. A reader may propose the simpler solution of looking for small dark blobs inside of a slightly lighter head. This solution works well under standard conditions, but images rarely have standard conditions. The lighting of the well changes the relative intensity of the eyes compared to the head. Some planaria have darker bodies than others. Sometimes specs of dust land on the head. A researcher can tell the difference between a spec of dust and an eye, but a simple blog detection algorithm cannot. Instead specifying what constitutes and eye under all of these varying condition by hand, by training a neural network it learns all of these rules itself. This justifies the additional computational costs and hand labeling costs.

4 Procedure

4.1 Setup

We used a Google Cloud Compute virtual machine with 8 CPUs, 52 GB of memory, and 1 NVIDIA Tesla K80 GPU. Our software is written in python 3, and uses TensorFlow[4], Keras[5], NumPy, SciPy, and openCV3.

4.2 Cropping Images

Before anything else, we crop the images to focus on the head of the planarian. This reduces the dimensions of the images from (1280x1024) to (100x100). We developed a method for automatically finding and cropping the head of the planarian in previous work. When a planarian moves it leads with its head, so we use image subtraction to find and crop the leading motion.

4.3 Hand Label Training and Test Sets

For the training set we labeled 2000 individual images.

- 500 images of planaria with no eyes drawn from 14 videos.
- 500 images of planaria with one eye drawn from 19 videos.
- 500 images of planaria with two eyes drawn from 22 videos.
- 500 images of poorly cropped images.

We drew from more than 10 videos for each category because many of the frames were poorly cropped or the worm was hidden from view.

We made two test sets. The first test set measures our accuracy on classifying individual images and the second test set measures our performance on classifying videos.

- The test set of individual images has 300 images from each category.
- The test set of videos has 16 videos of 0 eyed planaria 22 videos of 1 eyed planaria and 21 videos of 2 eyed planaria.

4.4 Data Augmentation

After our images have been labeled, make copies of each image rotated at different angles. By augmenting our training data we increase the overall size of our training data and make our network more robust to variations in input images.

4.5 Extract Features

We used the pretrained Inception V3 convolutional neural network from the TensorFlow GitHub to convert our cropped images into a set of features. We ran each cropped image from the training sets and the test set through the convolutional layers of Inception V3. We then flattened the outputs into 2048 vectors and saved the outputs. We used these saved values as the input to our own small densely connected neural network.

4.6 Training

After experimenting around with various architectures for our densely connected neural network we found that two hidden layers with 32 hidden units each performed the best. We used a stochastic gradient descent with a learning rate of 0.001, a momentum term of 0.9, and a weight decay of 0.001. We used 20% of our training set as a validation set, so while we trained for 30 epochs we only updated the network if it increased the accuracy on the validation set. We did this to prevent over-fitting to the training data.

4.7 Classifying Videos

Given a 50 frame video of a normal planarian with two, 10 of the frames of that video may view the planarian at a bad angle so it appears the planarian has only one eye. A poorly cropped frame may not see any of the planarian at all. Because of this many of the frames of a normal video are classified as abnormal or unknown by of neural network. As per the standards for classification, a video with a single frame displaying a normal two eyed planarian should be labeled as normal. Ideally we would classify videos as normal if our neural network confidently classified a single frame of a video as normal; however, our neural network produces too much noise to make decisions without a stronger signal. Instead we classify a video as normal if it confidently classifies at least 10 frames as normal with a confidence of over 85%. If our neural network confidently classifies less than 10 but more than 2 frames as normal, we classify the video as unknown.

5 Results

5.0.1 Accuracy On Individual Images

Our network achieved 76.6% accuracy on the test set of individually labeled images.

5.0.2 False Negatives

Our network achieved 0% False Negatives. We classified every video of an abnormal planarian as either abnormal or unknown.

5.0.3 False Postives

Our network also achieved 0% False Positives. We classified every video of a normal planarian as either normal or unknown.

5.0.4 Recall

Our network made predictions on 79.7% of the test videos. In other words, in classified 12 out of the 59 test videos as unknown.

5.0.5 Analysis

While our solution only makes predictions on 80% of videos, it has a very low error rate. For a toxin screening laboratory, having a low error rate is the top priority. Our solution has comparable error rates to the current method of labeling by hand, but takes much less time and is less expensive since it can be done on a computer rather than taking up the time of a researcher or lab assistant. A researcher or lab assistant will still have to classify the videos that our solution labeled as unknown, but hand-labeling 20%

of the data instead of 100% is still a major improvement. We also expect our recall to improve as we improve our methods.

6 Conclusion

The field of toxicology currently relies on expensive mammalian studies. Because of this toxin screening laboratories can not test all the environmental toxins that need testing. Planaria offer a low cost high throughput alternative. Researchers can breed and maintain large quantities of planaria at a low cost. They can also run many test simultaneously without the need for additional resources because automated test can reduce the work of researchers and lab assistants. We have developed a software system for automatically analyzing thousands of images that would otherwise have to be labeled by hand. When tested on the specific task of classifying the development of a planarian's eye, the system achieved error rates comparable to hand labeling. In the future we intend to use the same platform to classify the body shape of a planarian as normal or effected by the toxin. We also intend to increase the size of our training sets so as to improve our recall.

References

[1] Hagstrom, Danielle, et al. "Freshwater planarians as an alternative animal model for neurotoxicology." Toxicological Sciences (2015): kfv129.

[2] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

[3] Dai, Wenyuan, et al. "Translated learning: Transfer learning across different feature spaces." Advances in neural information processing systems. 2009.

[4] JJ Allaire and Dirk Eddelbuettel and Nick Golding and Yuan Tang. "tensorflow: R Interface to TensorFlow." https://github.com/rstudio/tensorflow (2016).

[5] Chollet, Francois, et al. "Keras." GitHub, https://github.com/fchollet/keras, (2015).

7 Acknowledgements

We would like to thank the Collins Lab for mentoring us and providing us with the planarian data used in this paper.