Digital Image Processing COT 5930

Dr. Scott Hammond

James Goedmakers

4/18/2024

https://github.com/jgoedmakers/COT5930FoveationProject

Foveation Techniques in Medical Image Analysis

Original

Foveated

Introduction

In the natural world, most animals have evolved a foveated visual system, in which their eyes perceive images at a higher quality near a central point of focus (fovea) than the area outside this area of focus, which is known as peripheral vision. In humans this adaptation is also present and readily observable, with image detail significantly higher at the location that our eyes are focusing, to a point that most people will be unable to accurately read text in their peripheral vision, but will have no issues once this same text is in their foveal vision.

Given the convergence toward foveated vision all across the biological world, it would be reasonable to assume that there are worthwhile advantages to this dynamic style of vision that would have encouraged this adaptation over a more static vision with quality consistent across the entire image. One of the more likely reasons for foveated vision is the potential improvements in efficiency in performing visual tasks. If only a small fraction of the field of view needs to be rendered in greater detail to accomplish the same visual tasks, then there would be substantial improvements in computational efficiency in foveated vision, which would reduce the energy demands placed on the organisms utilizing them. Given the historical success of foveated vision in biology, it may be possible to apply the same functionality to synthetic computer vision systems to attain the same benefits. In this report we will implement some digital foveation and apply it to a visual task of classification, in this case involving medical imagery, in order to evaluate possible benefits to incorporating foveated computer vision systems.

Experiment

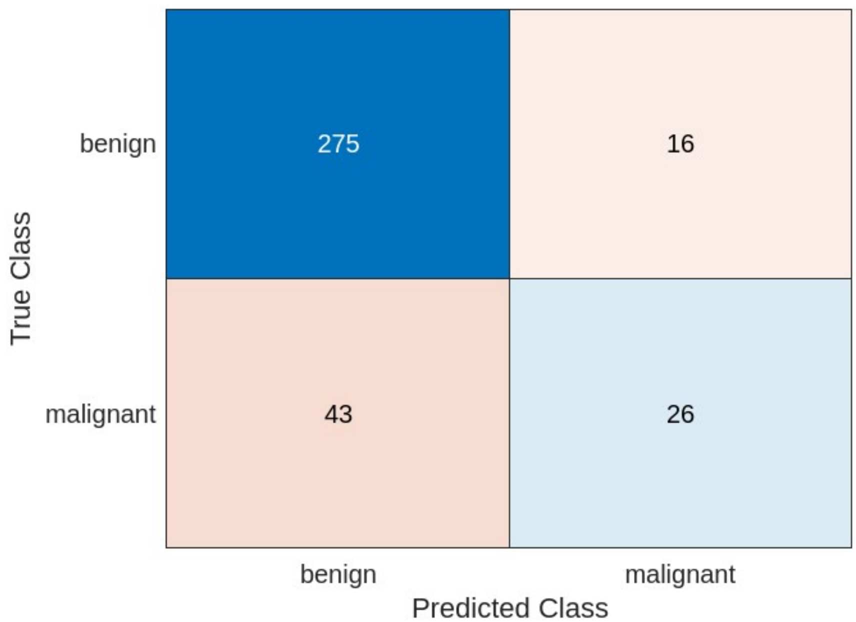
In our experiment, we will be performing image classification in a MATLAB environment using a deep convolutional neural network modified from GoogleNet. The dataset we will be utilizing is the ISIC 2016 dataset, which is a compilation of 900 malignant and benign skin lesions. The classification task will be to determine whether the skin lesion in each image are benign or malignant.

We will evaluate this classification task on the original ISIC dataset and a foveated version of the same dataset and compare the classification accuracy of the original and foveated images.

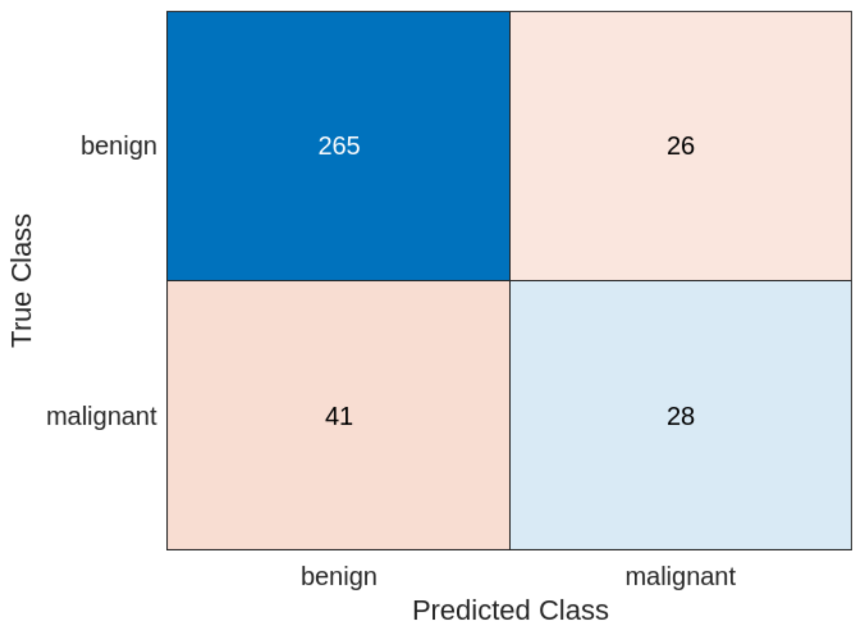
Results

Initial results show skin lesion classification at 83.61% accuracy for the original ISIC dataset and 81.39% for the foveated dataset.

Original Dataset



Foveated Dataset



One particular detail of note between the original and foveated tests is that most of the inaccuracy difference is found in 10 extra false positives, with only 2 additional false negatives. In medicine it is typically more tolerable to have false positives over false negatives, since false positives tend to lead to further testing to achieve confirmation while false negatives can lead to diseases going untreated.

Conclusions

The results of the experiment show some promise for foveated computer vision tasks. The classification accuracy difference between the original and foveated image datasets was only 2.22%, with the original achieving 83.61% and the foveated dataset achieving 81.39%. Although the current tests lack definitive efficiency metrics, we can estimate that there are reasonable efficiency gains to be made with minimal performance loss, given that the peripheral area of the foveated ISIC images was given 0.5 scaling.

Despite the appealing results of the experiment, there are multiple limitations in this test that should be addressed. The scope of the experiment is fairly limited due to the dataset used and binary classification task evaluated. As seen in some of the skin lesion dataset images, most images already have the lesion in question centered in the image, and are fairly circular in shape. This nature of this dataset is more ideal for foveated image processing which may have minimized performance drawbacks compared to images with irregularly shaped objects or objects that are out of center. This limitation can be alleviated with techniques such as identifying saliency points of an image to determine the optimal fovea region prior to classification. Another limitation is a lack of adequate comparison for compression and image size. In an ideal follow up experiment, we would be able to compare the classification performance of the original image to that of a compressed image of a smaller size, and then compare the compressed image to a foveated version of the image with identical size. In this way we could determine how much more effective foveated compression is in preserving classification performance compared to traditional compression methods.

References

Isic Challenge. (n.d.). https://challenge.isic-archive.com/landing/2016/39/

Floren, A., & Bovik, A. C. (2014). Foveated image and video processing and search. *Academic Press Library in Signal Processing: Volume 4 - Image, Video Processing and Analysis, Hardware, Audio, Acoustic and Speech Processing*, 349–401. https://doi.org/10.1016/b978-0-12-396501-1.00014-5

Jonnalagadda, A., & Eckstein, M. (2022). A foveated vision-transformer model for scene classification. *Journal of Vision*, *22*(14), 4440. https://doi.org/10.1167/jov.22.14.4440

Lukanov, H., König, P., & Pipa, G. (2021). Biologically inspired deep learning model for efficient foveal-peripheral vision. *Frontiers in Computational Neuroscience*, *15*. https://doi.org/10.3389/fncom.2021.746204