Imaging and Applied Optics 2018 (3D, AIO, AO, COSI, DH, IS, LACSEA, LSC, MATH, pcAOP) © OSA 2018

Classification of Digital Holograms with Deep Learning and Hand-Crafted Features

Tomi Pitkäaho[a], Aki Manninen,[b] and Thomas J. Naughton[a]

[a] *Department of Computer Science, Maynooth University, Maynooth, County Kildare, Ireland*

[b] *Biocenter Oulu, Faculty of Biochemistry and Molecular Medicine, University of Oulu, Finland*

tomi.pitkaaho@cs.nuim.ie, aki.manninen@oulu.fi, tomn@cs.nuim.ie

Abstract: Digital holographic microscopy allows a single-shot label-free imaging of living microscopic objects using a low intensity laser. Using reconstructed quantitative phase as an input to a convolutional neural network, detection of tumorigenic samples is possible.

OCIS codes: 090.1995, 100.6890, 170.6900

1. Introduction

Digital holographic microscopy (DHM) enables one to capture the full wavefront of a microscopic object [1]. DHM can be characterized as a label-free, low intensity, single-shot imaging technique that allows reconstructed phase to be used for microscopic object inspection and analysis. A magnified digital hologram $H_0(x,y) = |R|^2 + |O|^2 + R^*O + RO^*$ can be propagated at any depth z of the reconstruction volume with the Fresnel approximation. From the complex-valued reconstruction, both the phase and amplitude components can be extracted.

In this paper, we investigate the classification performance of a convolutional neural network (CNN) that is augmented by adding to the network layers that incorporate hand-crafted features. Classification results of healthy and tumorigenic multicellular samples with a multilayer perceptron (MP), CNN, and this new extended CNN (eCNN) are presented.

2. Convolutional neural network

Deep learning is an approach that enables simultaneous analyses at multiple levels of abstraction and can be used efficiently in different applications [2]. Deep CNNs have been used successfully in various visual object recognition and object detection applications [3, 4].

The proposed eCNN network architecture is shown in Fig. 1. The input to the convolutional part of the network is a 227×227 pixel unwrapped phase image. The network consists of five convolutional blocks followed by three fully-connected layers. Each convolutional block consists of convolution, batch normalization, and activation layers, followed by a max pooling layer. Hand-crafted and convolutionally extracted features are combined in the concatenation layer that is followed by a batch normalization, activation, and the last fully-connected layer. Each fully-connected block consists of fully-connected, batch normalization, and activation layers that are followed by a drop-out layer with 50% drop rate, with the exception of the two last fully connected layers that do not contain batch normalization, activation, or drop-out layers. The last layer in the network is a softmax layer that outputs class probabilities.

For training, validation, and testing of our approach we captured holographic data of healthy and tumorigenic multicellular Madin-Darby canine kidney (MDCK) cell clusters using an off-axis Mach-Zehnder digital holographic microscope with a 660 nm laser source and a 40X microscope objective. As a model for tumorigenic cells, we used MDCK cells that express an oncogenic G12V-mutant of K-Ras proto-oncogene ($\mathrm{KRas}^{\mathrm{V12}}$). The samples were incubated for six days before imaging. The set of holograms (2451 in total; 1972 healthy and 479 tumorigenic) was partitioned for training, validation, and testing. Due to the disbalanced dataset sizes, loss values (negative log likelihood loss) were multiplied by class specific weights that were 0.001 and 0.005 for healthy and tumorigenic samples, respectively. Batch sizes were 40 for the CNN and eCNN, and 200 for the MP. Each network was trained for 100 epochs by using an adaptive moment estimation (Adam) optimizer within the PyTorch framework.

3. Results

Figure 2 shows losses together with accuracies for each of the three approaches. Average test accuracies for three runs were 72.5%, 80.8%, and 87.5% for the MP, CNN, and eCNN, respectively.

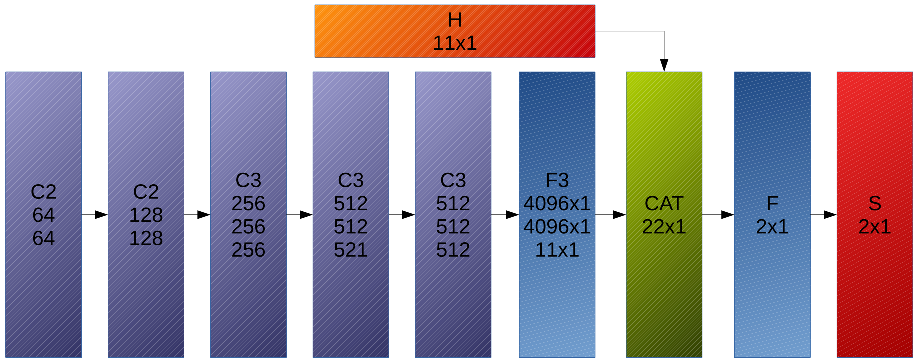

Fig. 1. Extended CNN architecture. C; convolutional block, F; fully-connected block, H; hand-crafted feature vector, CAT; concatenation block, S; softmax layer. If a block contains multiple layers, the number after block the definition shows the number of layers in a block; numbers below show the output depth from each layer. H, F, CAT, and S each output a one-dimensional vector.

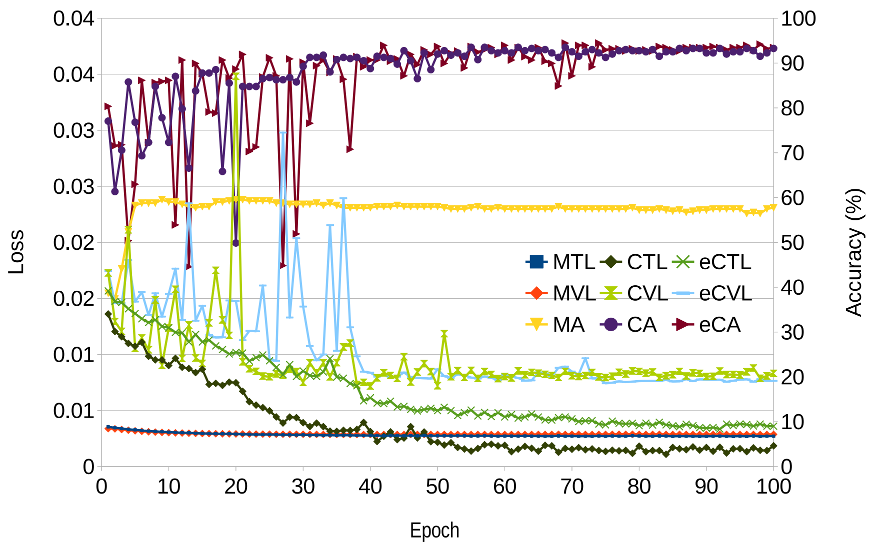

Fig. 2. Training loss, validation loss, and accuracy for each classifier, for a data set for which hand-crafted features proved beneficial for learning: MTL, MP training loss; MVL, MP validation loss; MA, MP validation accuracy; CTL, CNN training loss; CVL, CNN validation loss; CA, CNN accuracy; eCTL, eCNN training loss; eCVL, eCNN validation loss; eCA, eCNN accuracy.

Acknowledgements This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number 13/CDA/222, and an Irish Research Council Postgraduate Scholarship.

References

1. E. Cuche, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrast imaging," Optics Letters **24**, 291–293 (1999).
2. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).
3. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 1097–1105 (2012).
4. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv preprint arXiv:1512.03385 (2015).