

CLASSIFICATION OF CELL BIOMATERIAL INTERACTION TOXICITY LEVEL USING CONVOLUTIONAL NEURAL NETWORKS

Xhoena Polisi, Edit Dollani, Arban Uka

Computer Engineering Department, Epoka University

Tirana, Albania

xpolisi@epoka.edu.al

edollani16@epoka.edu.al

auka@epoka.edu.al

Keywords: cell, biomaterial, toxicity level, convolutional neural network

Abstract: *The medical image analysis field is highly dependent on good quality research time offering a noninvasive analysis and diagnosis of a medical problem. The most recent research credits the developments of different machine learning techniques for higher accuracy compared to traditional methods when analyzing more complex problems. Lots of research is being done towards the interactivity of the cells with different biomaterials, to provide an automatic model to determine the toxicity level. In this research, we will focus on different classification techniques using a convolutional neural network (CNN) of cells residing in different biomaterials. The problem is studied as multiclass classification: healthy, unhealthy, and severely disintegrated. The images are taken using a brightfield microscope, hence several image preprocessing techniques are applied to achieve high accuracy levels of predictions. The dataset used in this study includes approximately 20000 images for training and two different datasets with each more than 8000 images of size 128x128 pixels. LeNet architecture is used to analyze and classify the images. The network has four convolutional layers with kernels of size 5x5 applied, followed by max-pooling layers with the purpose of decreasing the number of weights. Without the preprocessing steps, the highest accuracy reached was 95%, whereas using several image preprocessing techniques increased the accuracy significantly. The highest resulting accuracy after preprocessing only the healthy part of the dataset with the Sobel filter was that of 99%.*

1 INTRODUCTION

Image processing and machine learning have been of major help in improving healthcare in terms of making predictions with accuracy and reducing a huge amount of tedious work that would require many human resources. Recently, deep learning and especially convolutional neural network is emerging as a prime machine learning method in computer vision. Deep Learning is a particular type of supervised learning whose base is artificial neural networks, inspired by the structure of biological nervous systems. Every artificial neuron in a network receives several inputs, computes the sum of those inputs, passes the result through a nonlinear function, and uses this result as an input to other artificial neurons. Over the last few years, several convolutional neural network architectures have been presented and achieved significant improvement in results. Contributions such as LeNet-5, AlexNet and other CNN architectures were proven to be efficient in various tasks including image recognition and image classification [1]

Studying the previous research, we see that other researchers have used LeNet Convolutional Neural Network (CNN) to classify two different types of bacteria. [2] LeNet which was originally presented by LeCun et al. in their 1998 paper, Gradient-Based Learning Applied to Document Recognition entails two sets of convolutional, activation, and pooling layers, trailed by a fully connected layer, activation, additional fully

connected, and lastly a SoftMax classifier. Other researchers have used multi-class classification using Support Vector Machine (SVM) with the same purpose of classifying bacteria sample images and managed to reach an accuracy of 97%. [3] SVM is a multiclass linear supervised technique used for classification especially in cases where the number of dimensions is bigger than the number of samples.

The LeNet-5 convolutional model was also used as a base model for other convolution neural network models. An example of that is in paper [4] where the authors added an extra convolutional layer along with a pooling layer to deepen the network. Another difference made to this model was the connection of the backward propagation of the first two pooling layers to the last pooling layer through convolution. The purpose of this modification is to make the most of the low-level features extracted by the network. The experimental results of the classification were good.

In [5] the authors propose the idea of having a CNN multi-class classification framework used for dermoscopy samples. They used as an upper hand the fact that those multiple classes can be later united into two classes and turn the classification into a binary one. The CNN used in [5] is GoogLeNet Inception-v3 and the results on the multi-class classification show an impressive increase of accuracy of 7%.

In [6] the authors have reviewed Hep-2 cell classification using different deep learning methods and compared them based on performance. The classification as the authors of the paper described is done at two extents, the cell, and the specimen. Furthermore, they also compare existing datasets and look at the possibilities for future research.

In different papers, cells have been classified using different features, such as color, geometric and texture. Color feature is generally connected to visual appearance of the cells. In focus are taken characteristics such as hue, saturation, and brightness. A very important technique to consider for this is the histogram of the sample which shows a graphical representation of the number of pixels as a function of their intensity. Geometric features are best described by different characteristics such as area and perimeter of cells, shape of nucleus or cytoplasm in cells and other details pointed out by medical experts. Textures feature is focused on patterns of material, color or intensity that can be visually detected. [7]

The authors of [8] present an Optimal Feature Selection for medical images using deep learning with the purpose of bringing more attention to feature selection and classification. The aim of their work to attain an optimal feature selection classification was successfully reached with an accuracy of 95.22%.

An important role in the medical field is the detection, counting and classification of various types of cells. However, this work can be very difficult considering the variety of the biological variability and the limitations in quality of cell samples. [9]

As mentioned in [10] the analysis of the cell images can be affected from different conditions such as non-uniform illumination, gray shades because of low contrast, translucency of the cytoplasm of the cell. Another issue is that deep learning algorithms require many samples which is a restriction for any researcher since the medical labor to get those samples is a tedious and time-consuming job.

High performance computing based on GPU is crucial for image processing in the medical field, as it is needed to reduce time when training deep learning models. [5, 10] This is because the most important aspects image processing is based on are image size, speed, and resolution. GPU has data processing capacity that exceeds that of CPU and makes it easier to work on high-performance computing on ordinary computers. [11]

It is recognized the fact that cell datasets are small sized since they are difficult to obtain and considering deep learning usually requires many samples the problem in overfitting may arise. A solution for this in the absence of the opportunity to have a larger dataset is using data augmentation approaches. These techniques which are simple and effective are more commonly cropping, rotating, and flipping of sample images. Except for data augmentation, other solutions could be batch normalizations and dropout. Although having a larger dataset is expensive in computation time, it accelerates deep learning in terms of development.

In our research we will focus on classifying cells based on their health state: healthy, unhealthy, and

severely disintegrated. The focus will firstly be on preprocessing the samples. The paper is organized in the following order. Chapter 2 talks about the materials and dataset used, whereas the method is described in chapter 3. Results are shown on chapter 4 and concluded on chapter 5.

2 MATERIALS

2.1 Challenges of the dataset

The dataset is obtained using brightfield microscopy with no cell staining when they are being in contact with various biomaterials. A main challenge is dealing with the difficulties of unstained images since more often they suffer from nonuniform illumination, low contrast, and transparency of the cytoplasm (see Figure 1).

To this challenge can be added the fact that the original dataset is small compared to the datasets needed to train CNNs. Furthermore, the CNN model should be adapted to deal with the difficulties of unstained images as mentioned above and should be trained carefully to make reliable predictions. [10]

Other than that, the samples show cells of different sizes and structures which can be confusing for the

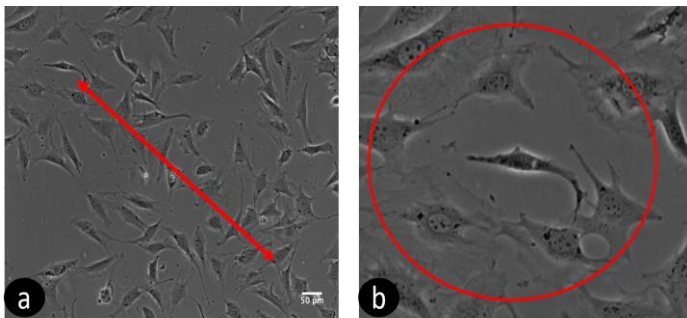


Figure 1 Some of the challenges of the dataset: a) the non uniform illumination and b) the low contrast and the transparency of the cytoplasm

training of the deep learning model. They are elongated, star shape alike while residing in healthy environment, whereas a more circular and shrinking appears in toxic environments. When the material is highly toxic, we may not find cells at all, on cell debris (refer to Figure 2). These microscopy cell samples are crucial in the medical field to initially decide the health state of the cell and later determine the efficacy of various treatments. To overcome these challenges several preprocessing techniques with the purpose of achieving high accuracy in multi class classification are applied and artificial dataset augmentation is used for training the network.

2.1 Dataset information

The dataset used in this study is divided into other datasets, this way fulfilling the need of having different datasets to use for training and testing the classification task. The main dataset used for training has sample images of size 128 x 128 pixels. The classification will be done for three classes: Healthy, Unhealthy and Severely Disintegrated. The Severely Disintegrated class represents cells that are in a bad condition, disintegrated cells. The training earlier in the research was done using this dataset with the purpose of the training being done in a shorter amount of time and to lessen the complexity for the neural network model since the images are only 128x128 pixels. In most of the 128 x 128 samples there should be at least one cell. Some samples which contained no cells were removed

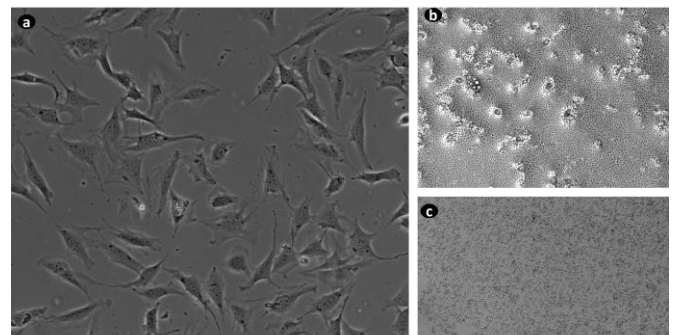


Figure 2 Cells change their shape and size while residing in healthy (a), unhealthy (b) and highly toxic (c) environment

because later during the training they can reduce the accuracy obtained

To measure the performance of the trained model larger images of size 1280 x 1024 pixels were cropped into 80 sample images of size 128 x 128 pixels and these crops were tested for accuracy. Experiments were done with different datasets with the goal of seeing the difference in the performance of the model. The ratio of the division was kept 80% for training and 20% for testing.

3 METHODOLOGY

The Convolutional Neural Network chosen for multiclass classification using the mentioned dataset is LeNet, as one of the most used architectures based on the literature. LeNet architectures follows a model called Sequential model. This model is like a stack of layers where each layer has one input and one output. This model is very feasible to be implemented as it allows you to easily build a model adding layers one by one. In order to overcome dataset challenges, different preprocessing techniques will be used

3.2 Preprocessing

Different techniques can be used for preprocessing images such as grayscale, histogram equalization, etc. [3, 12] Every image preprocessing technique has its own advantages and disadvantages. The effectiveness of an image processing technique also varies from the images of the dataset to be processed and their properties. Different methods of preprocessing have different resolution or noise levels. Filters that can be used for image processing can be grouped as nonlinear and linear. [13] Nonlinear are the filters whose output is not a linear function of its input. Examples of nonlinear filters are median filter, bilateral filter, etc. When using Linear filters, the value of output pixels is expressed as a linear combination of the values of the neighbors of the input pixel. A disadvantage of linear filters is the risk that since they act as low pass filters, linear filters can smoothen the edges rather than enhancing them and amplify the noise. Examples of linear filters are Gaussian filters.

In previous studies as image preprocessing techniques on cell samples have been used: soft clustering using Gaussian mixture models, various color components, watershed transform etc. Watershed transform application permits the locating of 3 regions of interest: the nucleus of the cell, the entire cell and the area surrounding it. [7] Often histogram equalization is exploited to enhance the quality of the input sample. [8]

Edge detection is also a very used and crucial technique in image processing. However, it is often difficult to use in medical images because of the sample conditions which vary to the exposure to other aspects. The Sobel filter is an image processing filter which is used for edge detection in images by emphasizing the edges.

Several testing along with different preprocessing techniques were before deciding which one will be used with the trained model, to reach the goal of getting the higher accuracy in classification. The first one we can mention is Sobel filtering, with two different kernels, one for each direction. In image processing a kernel, alternatively called a mask or convolution matrix is a matrix used for purposes of blurring, sharpening edges, smoothening images etc. The second one is Laplacian filtering with a window size of 5x5. Bilateral Filter is a noise-reducing, non-linear filter which preserves the edges and smoothenes the images. It works similarly to the Gaussian Filter by replacing the pixel intensities with the average of the intensity values from the nearby pixels. The median filter as mentioned previously in the literature review is a non-linear filter used with the purpose of reducing noise from an image. The median filter replaces the gray level of each pixel with the median of the gray levels of the pixels surrounding the input pixel.

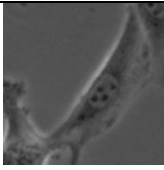

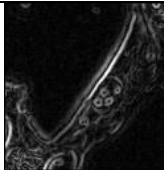
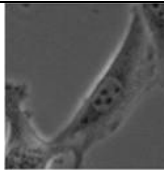
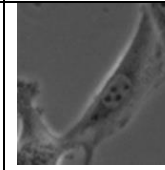
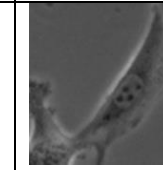
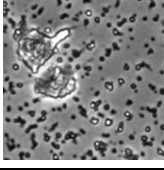
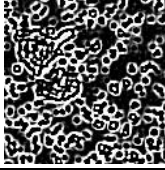
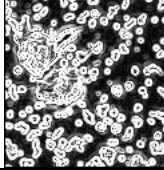
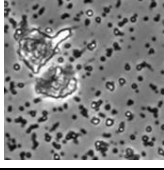
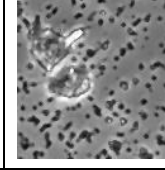

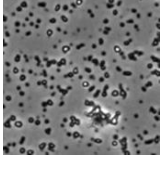

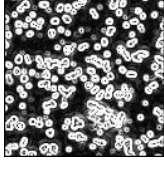
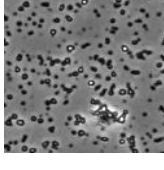
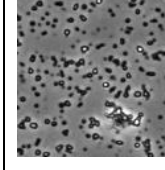
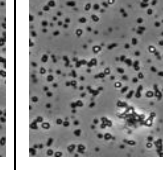
A wavelet denoising filter represents the wavelet representation of the image in which the noise is represented by small values in the wavelet that are set to 0.

The VisuShrink method employs only one universal threshold to all wavelet coefficients which removes

Gaussian noise with high probability and functions by smoothening the sample appearance. BayesShrink is another method used for wavelet denoising where a unique threshold is estimated for each wavelet subband which is considered better than what can be obtained with only one threshold.

Based on the results, (some of them shown in Table 1), the preprocessing method decided for this work are Laplacian and Sobel filtering. The other filters did not show any meaningful change on the preprocessed image compared to original.

Table 1 Different Preprocessing techniques applied on cells residing in healthy and toxic environment

Environment	Preprocessing techniques					
	None: Original	Laplacian Window size=5	Sobel	Bilateral Filter	Median Filter	VisuShrink & BayesShrink
Healthy						
Unhealthy						
Severely disintegrated						

3.3 Neural Network Architecture

The architecture used was the LeNet architecture. Starting with the input layer, the images are resized to be of size 128 x 128 pixels. The network has four convolutional layers with kernels of size 5x5 applied, followed by maxpooling layers with the purpose of decreasing the number of weights. The four sets of convolutional layers are like CONV => RELU => POOL. Dropout is applied to avoid overfitting.

4 RESULTS

4.1 Parameter Definition

The classification initially is done using two classes: healthy, unhealthy based on the environment where the cell resides. Another class is added after noticing a meaningful change among unhealthy environments, which disintegrated the cells residing there completely. This is crucial as no live cells can be detected later in these images, when using segmentation techniques.

To check the accuracy of our method, different metrics will be used such as precision, recall, F1 score etc. Precision is an important measure to determine how precise the model is by representing the ratio of the correct predictions of the positive to the total predictions of the positive. High precision indicates a low false

positive which is good. The Recall metric represents how many of the true positives are predicted as positive from the model. F1 score is a metric which is used to seek for balance between Precision and Recall so it uses both false positives and false negatives. In cases of uneven class distribution, it is a better indicator than accuracy. The performance of the classification model can be displayed using a receiver operating characteristic curve also known as ROC curve which represents True Positive Rate vs False Positive Rate. The TCR vs FPR are shown plotted in different classification thresholds. A lower classification threshold would indicate an increase in both TP and FP, classifying more objects as positive. AUC is the area under the ROC curve. The higher the AUC score, the better the classifier performs in a classification task.

We will compare our results on different datasets, changing the batch size, the preprocessing method, and the number of images used for training and testing.

The batch size is a hyperparameter that represents the number of samples to work with before revising the inner model parameters. It works similar to a loop iterating over the samples and predicting results which when the batch is finished are compared to the expected output. Furthermore, an error is calculated which is used as an update algorithm for the model. Most used mini-batch gradient descent which we are also going to use and test are: 32, 64 and 128.

The first experiment is done by changing the batch size to 32 and 64, and leaving the number of epochs unchanged, and no preprocessing step is performed. Training was done on a dataset (12520 Healthy (noted as H) + 7582 Unhealthy (noted as U) + 704 Severely Disintegrated (noted as N)). The testing was done on two different datasets: i) Dataset 1 (5607 H + 2641 U + 176 N) and ii) Dataset 2 (5617 H+5464 U + 176 N)

The second experiment is done by changing the epoch number in order to find the best parameters for testing. Based on the results shown below on Figure 3, the best parameters to be used are number of epochs 60 and batch size 32. As a base overall accuracy of training for the next experiments to be compared to will be used that of 95% reached by the experiments.

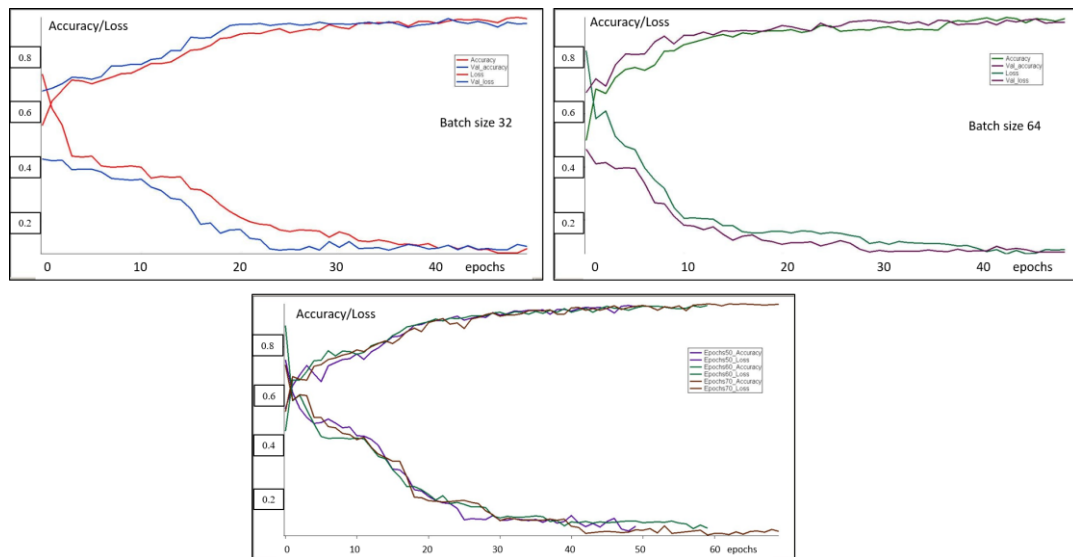


Figure 3 Graphs of accuracy/loss for batch size 32 (upper left) and 62 (upper right), and for different number of epochs (down)

4.2 Classification with preprocessing

The preprocessing techniques will be tested on two models with certain parameters called model A and B. Model A has a batch size of 32 and number of epochs 60 and model B has a batch size of 64 and

number of epochs 45. Model A performed rather well with a training accuracy of 95%. Testing on the Dataset 1 resulted in an accuracy of 94%, whereas on Dataset 2 the acquired accuracy was 79%. This model will serve as our initial results and will be used for comparison of the other models and experiments. The dataset used for testing was chosen Dataset 1, as a higher accuracy compared to Dataset 2. Testing of data using Laplacian as a preprocessing method results in a model accuracy of 96% so we have an increase in the accuracy compared to the base model by 1%. Sobel filtering is used as a preprocessing method only on the images labeled as healthy, since it was noticed a decrease on the accuracy when applied for the whole dataset. Classification results after using this preprocessing method have a 98% accuracy, which indicates that the performance of the model is greatly improved. A test was run also on Dataset 2, and the accuracy was still 99%. This indicates that usage of Sobel filtering as a preprocessing method increases the performance of the model significantly. When Model B is used, the overall training accuracy is 96%, whereas when tested on Dataset 1 the accuracy is 95%, and on Dataset 2 is 79%. When preprocessing is used the accuracy is lower for Laplacian compared to Model A (95%), but increases when Sobel Filtering is used (99% for Dataset 1 and 99.9% for Dataset 2). ROC graphs of these experiments are shown on Figure 4. Overall we can conclude that the usage of Sobel Filtering as a preprocessing technique increases the accuracy of cell classification, and Table 2 shows the results of its performance for every metric.

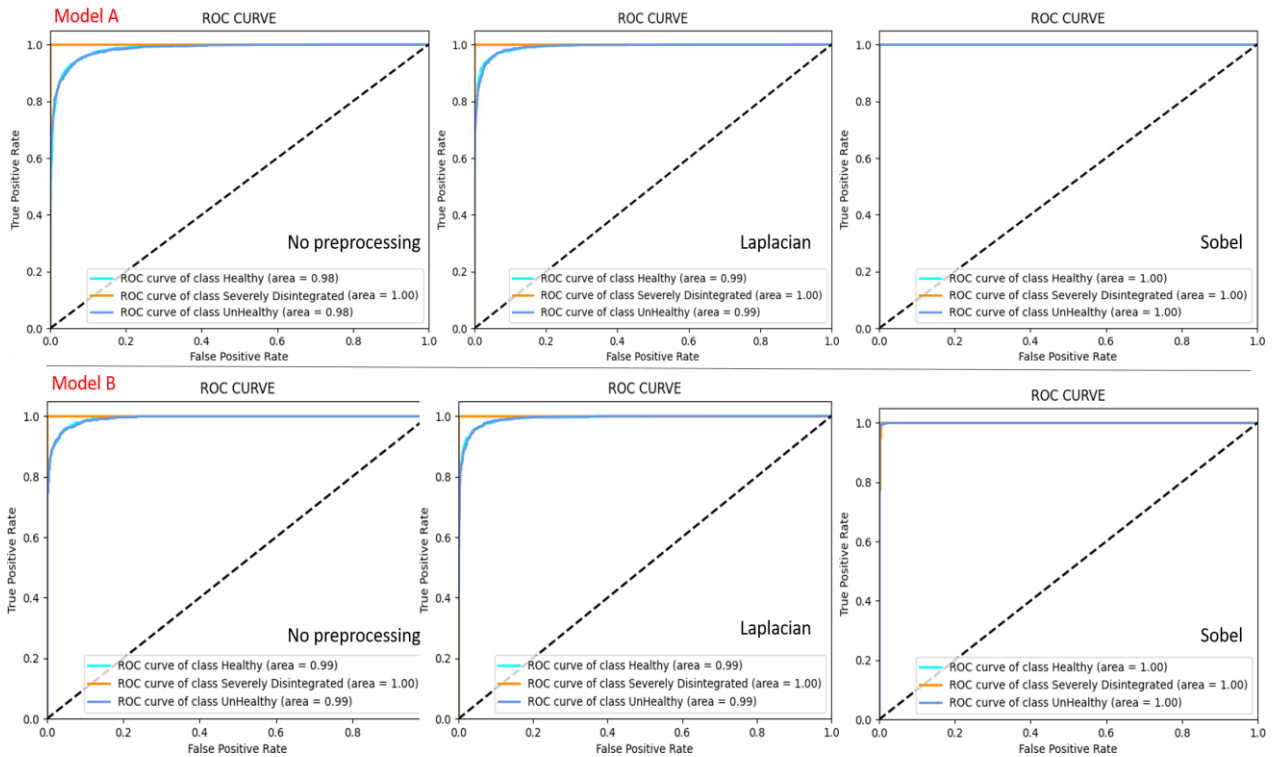


Figure 4 Graphs of ROC for Model A and B using no preprocessing of the data, Laplacian and Sobel filter as preprocessing methods

Table 2 Results of the metrics used to compare Model A and B when Sobel Filtering is used as a preprocessing method

	Precision		Recall		F1-score		Support	
	Model A	Model B	Model A	Model B	Model A	Model B	Model A	Model B
Healthy	1.00	1.00	1.00	1.00	1.00	1.00	2504	2504
Severely Disintegrated	0.64	0.77	0.99	1.00	0.98	0.87	141	141
Unhealthy	1.00	1.00	0.95	0.97	0.97	0.99	1517	1517
Accuracy					0.98	0.99	4162	4162
Macro avg	0.88	0.92	0.98	0.99	0.92	0.95	4162	4162
Weighted avg	0.99	0.99	0.98	0.99	0.98	0.99	4162	4162

5 CONCLUSIONS

The accuracy of medical image analysis relies heavily on the availability of large datasets to be analyzed and on the results to be predicted in a short time. The medical image classification field is one that is in need for large amounts of sample data to be analyzed and predicted in a short time. This research was focused on the combination of convolutional neural networks with several preprocessing techniques with the purpose of obtaining high classification accuracy. The study relies on multiclass classification to predict three different cell health levels. The model was trained on a dataset with more than 20000 images and tested on two different datasets with each more than 8000 images. While without preprocessing the dataset, the highest accuracy reached was 95%, with the dataset preprocessed using several methods, there was a clear improvement in the classification accuracy of the model. The highest resulting accuracy after preprocessing only the healthy part of the dataset with the Sobel filter was that of 99%.

REFERENCES

- [1] A. N. H. A. P. X. T. A. O. I. G. & V. Uka, "Basis of Image Analysis for Evaluating Cell Biomaterial Interaction Using Brightfield Microscopy.," in *Cells, tissues, organs*, 210, 2021
- [2] B. A. a. A. H. M. Mohamed, "Automated classification of Bacterial Images extracted from Digital Microscope via Bag of Words Model," in *2018 9th Cairo International Biomedical Engineering Conference (CIBEC)*, 2018.
- [3] T. a. K. J. H. a. B. H. a. L. S. J. Go, "Machine learning-based in-line holographic sensing of unstained malaria-infected red blood cells," *Journal of Biophotonics*, vol. 11, 2018.
- [4] T. a. P. S. Treebupachatsakul, "Bacteria Classification using Image Processing and Deep learning," in *2019 34th International Technical Conference on Circuits/ Systems, Computers and Communications (ITC-CSCC)*, 2019.
- [5] E. a. W. X. a. W. C. a. D. Y. Chen, "Application of Improved Convolutional Neural Networking Image Classification," in *2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, 2019.
- [6] N. a. L. E. a. T. K. a. S. H. Meng, "Large-Scale Multi-Class Image-Based Cell Classification With Deep Learning," *IEEE Journal of Biomedical and Health Informatics*, 2018.
- [7] W.-C. L. a. J.-W. Wang, "Edge detection in medical images with quasi high-pass filter based on local statistics," *Biomed. Signal Process. Control.*, vol. 39, pp. 294-302, 2018.
- [8] I. a. I. R. a. C. A. a. D. R. a. B. R. Mocan, "Automatic Detection of Tumor Cells in Microscopic Images of Unstained Blood using Convolutional Neural Networks," in *2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP)*, 2018.
- [9] A. a. P. X. a. H. A. a. D. C. a. V. N. E. Uka, "Analysis of cell behavior on micropatterned surfaces by image processing algorithm," in *IEEE EUROCON 2017 - 17th International Conference on Smart Technologies*, 2017.
- [10] B. H. a. A. B. a. A. Hajdu, "Assisted deep learning framework for multi-class skin lesion classification considering a binary classification support," *Biomedical Signal Processing and Control*, vol. 62, p. 102041, 2020.
- [11] R. J. S. a. S. S. J. a. P. I. V. a. P. D. A. a. G. D. a. S. K. Raj, "Optimal Feature

Selection-Based Medical Image Classification Using Deep Learning Model in Internet of Medical Things," *IEEE Access*, vol. 8, pp. 58006-5801, 2020.

- [12] G. J. ., & Z. D. Cai L., "A review of the application of deep learning in medical image classification and segmentation," *Annals of translational medicine*, vol. 8, 2020.
- [13] M. S. a. M. D. a. B. S. a. J. N. a. J. N. a. C. J. Norouzzadeh, "A deep active learning system for species identification and counting in camera trap images," *Methods in Ecology and Evolution*, vol. 12, pp. 150-161, 2021.