

# Estrous Cycle Classification through Automatic Feature Extraction

Gerardo Hernández Hernández<sup>1</sup>, Leonardo Delgado Toral<sup>2</sup>, María del Rocío Ochoa Montiel<sup>1</sup>, Erik Zamora Gómez<sup>1</sup>, Humberto Sossa<sup>1,5</sup>, Aldrín Barreto Flores<sup>2</sup>, Francisco Ramos Collazo<sup>3</sup>, Rosalina Reyes Luna<sup>4</sup>

<sup>1</sup> Instituto Politécnico Nacional,  
Centro de Investigación en Computación,  
México

<sup>2</sup> Benemérita Universidad Autónoma de Puebla,  
Facultad en Ciencias de la Electrónica,  
México

<sup>3</sup> Benemérita Universidad Autónoma de Puebla,  
Bioterio Claude Bernard,  
México

<sup>4</sup> Benemérita Universidad Autónoma de Puebla,  
Facultad de Ciencias Biológicas,  
México

<sup>5</sup> Tecnológico de Monterrey,  
Campus Guadalajara,  
México

{gerardohernandez.hernandez, delgado.200917263, ma.rocio.ochoa, ezamora1981, humbertosossa}@gmail.com, abarreto@ece.buap.mx, francisco.collazo@correo.buap.mx, rreysluna@hotmail.com

**Abstract.** We study and propose, for the first time, an autonomous classification of the estrous cycle (the reproductive cycle in rats). This cycle consists of 4 stages: Proestrus, Estrus, Metestrus, and Diestrus. The short duration of the cycle in rats makes them an ideal model for research about changes that occur during the reproductive cycle. Classification is based on the cytology shown by vaginal smear. For this reason, we use manual and automatic feature extraction; these features are classified with support vector machines, multilayer perceptron networks and convolutional neural networks. A dataset of 412 images of the estrous cycle was used. It was divided into two sets. The first contains all four stages, the second contains two classes. The first class is formed by the stages Proestrus and Estrus and the second class is formed by the stages

Metestrus and Diestrus. The two sets were built to solve the main problems, the research of the reproductive cycle and the reproduction control of rodents. For the first set, we obtained 82% of validation accuracy and 98.38% of validation accuracy for the second set using convolutional neural networks. The results were validated through cross-validation and F1 metric.

**Keywords.** Estrous cycle, GLCM, machine learning, convolutional neural network, multilayer perceptron, SVM.

## 1 Introduction

The estrous cycle is the period between two ovulations and defines the receptivity of the female

for its reproduction. The short duration of the cycle in rats makes them an ideal model for investigating the changes that occur during the reproductive cycle [24, 17, 25, 7]. The cycle is used for reproduction control of rodents, obtaining high fertility rates, scheduling of production on specific dates, knowing the time of gestation, and the development of embryos at specific ages. The estrous cycle in rats lasts four days and is characterized by the stages: Proestrus, Estrus, Metestrus, and Diestrus, which can be determined according to the cell types observed in the vaginal smear [3].

The objective of this work is to provide a tool for automatic detection of estrous cycle stages through the image processing and neuronal networks. For this, we use manual and automatic feature extraction techniques, as well as different classifiers such as radial kernel Support Vector Machines (SVMs), Multilayer Perceptron networks (MLPs), and Convolutional Neural Networks (CNNs). The dataset, the feature extraction algorithms, and the neural network models are available in [10].

The rest of the paper is organized as follows. Section 2 provides a review of classical methods used for classifying the estrous cycle and shows the similarities between Papanicolaou (PAP) cells and cells of the estrous cycle. In Section 3 we review the feature extraction and the neural networks used. In Section 4 we show our approach for classifying the estrous cycle. We present several experiments and results, as well as a short discussion to demonstrate the effectiveness of our proposal. Finally, in Section 5 the conclusions and directions for future research are discussed.

## 2 Previous Work

Shannon L. Byers [3] describes a variety of methodologies used for classifying the estrous cycle. All of them need specialized people for being performed (Table 1). In the same paper, a graphic tool for classifying with the estrous cycle is presented (Figure 1).

Claudia Caligioni [4] describes the process to follow for classifying the estrous cycle through vaginal smear.

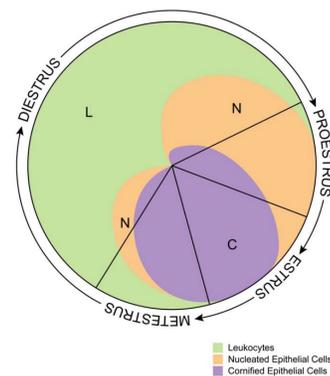


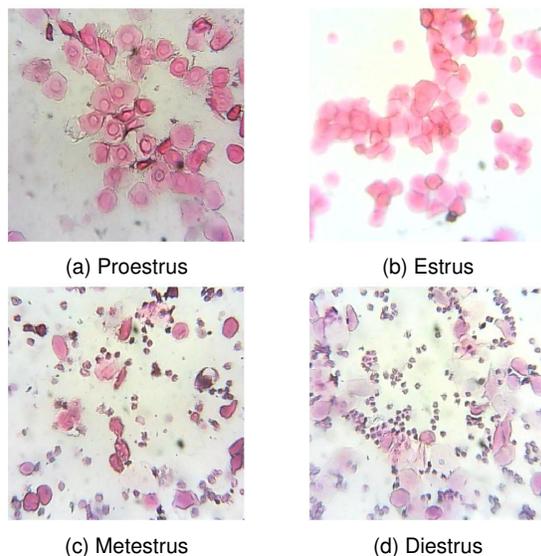
Fig. 1. Graphic tool presented by Byers on [3]

Her work shows that it is not necessary to dye the rats to classify the sample. According to her work and the work of Marcondes [18] the percentage of each kind of cell present in the vaginal smear must be:

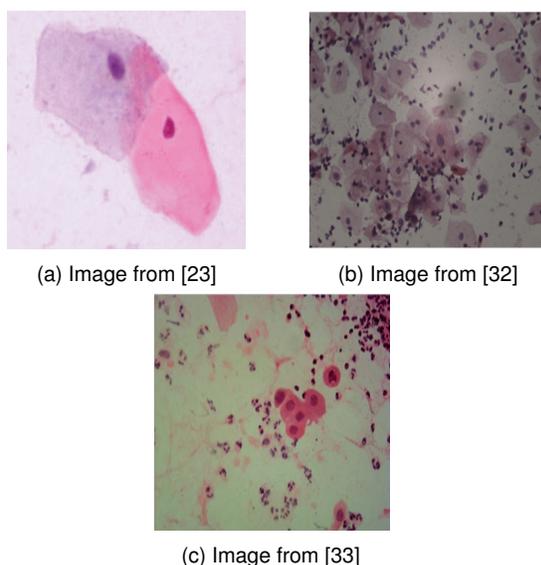
- **Proestrus:** Predominance of nucleated epithelial cells (Figure 2-a). These cells may appear in clusters or individually) [4].
- **Estrus:** Is characterized by cornified squamous epithelial cells, which occur in clusters (Figure 2-b). There is no visible nucleus; the cytoplasm is granular, and the shape is irregular [4].
- **Metestrus:** Is a mix of cell types with a predominance of leucocytes and a few nucleated epithelial and/or cornified squamous epithelial cells (Figure 2-c) [4].
- **Diestrus:** This stage consists predominantly of leucocytes (Figure 2-d) [4].

The state-of-the-art of estrous classification cycle does not have an automatic way for being performed at this moment. Nevertheless, the PAP cells are similar to the estrous cycle cells (Figure 3), and concerning this topic, there is a wide variety of algorithms.

Mariana E. Plissiti [22] has a 99.39% of accuracy segmenting the nucleus on PAP cells through the watershed transform. In the same year, she [21] worked on a methodology for segmenting the nucleus and overlapping cells through the



**Fig. 2.** Stages of the estrous cycle, images obtained by Vivarium Claude Bernard



**Fig. 3.** PAP Images

H transform, SVM and C-Means. In 2015 [23] she presented a method for the segmentation of cytoplasm that showed an overlapping. She uses a variety of intensities in the RGB obtained from super-pixels. Dabashree Kashyap [12] propose a method to classify PAP cells using geometric and texture characteristics. He uses GLCM metrics and 3 SVM where the best result is obtained using a polynomial kernel. Ling Zhang [32] presents a method for segmenting cells using graphs.

He takes an image in the CIELAB color space (channel A) and applies the three-threshold Otsu algorithm; finally, he uses a graph-cut approach for enhancing the segmentation. The results obtained are robust to non-bimodal distributions in the histograms of the image. Meng Zhao [33] proposes a method of classifying through image analysis by blocks slightly larger than the size of a cell. For the classification, he uses texture and histogram metrics. He reaches a 98.98% accuracy and 95% sensitivity. The detection of the estrous cycle keeps similarities with the classification of PAP cells.

However, the main difference is that the classification of the estrous cycle is based on the quantity of each kind of cell presented in the samples [4, 18] and the PAP classifying is based on the morphology of the cells. We took the PAP work as a starter point. Our proposal eliminates the bias of designing manual feature extractors, using automatic feature extraction, and achieving a completely automatic classification through neural networks of last generation.

### 3 Methodology

This section describes the different approaches used for the classification of the four aforementioned estrous cycles. As a first approximation, SVMs and MLPs are trained for the manual feature extraction method, described in Section 3.2. Subsequently, two architectures of Convolutional Neural Networks are trained, the first from scratch, the LeNet-5 network, and the second, a pre-trained model with Imagenet dataset [8], the VGG16 network.

**Table 1.** Advantages and disadvantages of the methodologies used for classifying the estrous cycle

Methodology	Advantages	Disadvantages
Cytology	Identifying all stages.	The sample needs to be dyed.
Impedance	The cycle can be measured using a probe.	The methodology only can classify one stage.
Observation	No extra equipment is required.	The methodology only can classify one stage.

### 3.1 Pre-Processed Features

The cell image is converted to the CIELAB color and grey level space. From CIELAB the channels A and B which represents a variation of the red color to green and the blue color to yellow are taken. A median filter with a window of 5,5 to eliminate the noise it is applied. Otsu's approach [20] is applied to segment the regions of interest.

### 3.2 Manual Feature Extraction

For the extraction of texture features, we used the GLCM algorithm [5, 9] with steps ( $\delta$ ) [1,5] and angles ( $\theta$ ) [ $0^\circ, 45^\circ, 90^\circ, 135^\circ$ ]. The GLCM metrics used:

Contrast:

$$O = \sum_{i,j=0}^{l-1} P_{i,j} (i - j)^2. \quad (1)$$

Dissimilarity:

$$D = \sum_{i,j=0}^{l-1} P_{i,j} |i - j|. \quad (2)$$

Homogeneity:

$$H = \sum_{i,j=0}^{l-1} \frac{P_{i,j}}{1 + (i - j)^2}. \quad (3)$$

Energy:

$$E = \sum_{i,j=0}^{l-1} P_{i,j}, \quad (4)$$

where  $P_{i,j}$  (the co-occurrence probability between grey levels  $i$  and  $j$ ) is defined as:

$$P_{i,j} = \frac{G_{i,j}}{\sum_{i,j=0}^{l-1} G_{i,j}}, \quad (5)$$

where  $G_{i,j}$  represents the number of occurrences of grey levels  $i$  and  $j$ , given a certain pair  $(\delta, \theta)$ . The matrix  $G$  is calculated considering the times in which occurrences of intensity change given an angle  $\theta$  and a step  $\delta$ . Each element of the matrix represents the number of occurrences between a change of intensity to another (Figure 4).

The morphological metrics obtained were the number of elements, the average compactness Equation 8 and the total compactness Equation 7 giving a total of 105 characteristics, 35 by each channel of the CIELAB and 35 by the greyscale.

Compactness factor:

$$FC = \frac{Area}{Perimeter^2}. \quad (6)$$

Total compactness:

$$CT = \sum_{i=1}^{Nl} FC_i. \quad (7)$$

Average compactness:

$$CP = \frac{\sum_{i=1}^{Nl} FC_i}{Nl}, \quad (8)$$

where  $Nl$  is the number of connected elements in the binary image.

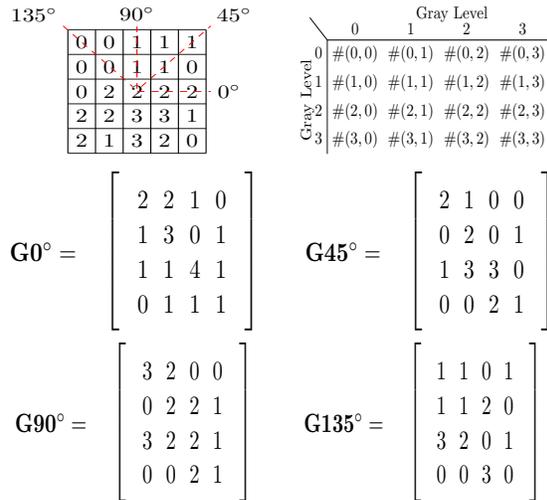


Fig. 4. a) The 5x5 image with four grey-levels, b)The general form of any GLCM for images with grey-levels values 0-3, c)-f) Calculation of GLCM for  $(\delta, \theta)=(1,0),(1,45),(1,90),(1,135)$

### 3.3 Classification

Nowadays, there is a great diversity of classifiers. These are divided into two large branches, supervised and unsupervised. Within the unsupervised classifiers we can find algorithms like k-Means, Self-organizing maps, Hidden Markov models.

Supervised algorithms have better acceptance in the scientific community because they are more effective methods to achieve higher classification performance. Within this other branch, we can find the Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), Deep Neural Networks (DNN) and the Convolutional Neural Network (CNN).

The nature of these classifiers is varied, while the SVMs excel in their ability to classify patterns (data), they have a low performance when trying to classify images. The MLPs and DNNs networks can classify both patterns and images of a single channel; however, these models tend to be large and difficult to train.

Finally, we have the CNNs, which in recent years, have proved to be the ideal tool for image classification with automatic feature extraction.

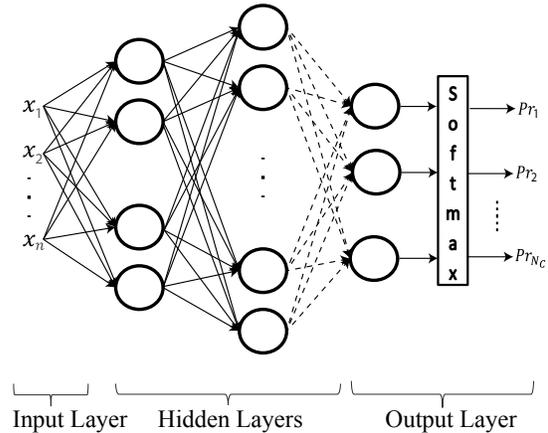


Fig. 5. DNN classic architecture.

### 3.3.1 Deep Neural Networks

Deep Neural Networks (DNNs) are composed of a computation unit called Perceptron, defined by Equation 9. The objective is to separate two classes through a hyperplane.

$$y = \sum_{i=1}^n x_i w_i + b, \tag{9}$$

$$O = F(y), \tag{10}$$

where  $x_i$  represents the  $i$ th element of the input  $x$  vector.  $w$  are the trainable synaptic weights,  $b$  is the bias and  $F$  represents a non-linear activation function [30, 27, 26]. In our case, the  $ReLU = \max(0, y)$  function [19].

A DNN network usually consists of more than three intermediate layers, each layer of a variable Perceptron number, as shown in Figure 5. These networks are trained by stochastic gradient descent.

### 3.3.2 Support Vector Machine (SVM)

A SVM is a discriminative classifier, whose objective is to look for a separation hyperplane between two classes, defined by Equation 11. This maximizes the separation distance between two classes. For this, the SVM uses non-linear transformation functions or kernels. The two most common kernel functions are: the linear ( $K_l$ ),

Equation 12, and the radial basis ( $K_g$ ), Equation 13, like Gaussians [6, 2].

$$D(x) = \sum_{k=1}^p \alpha_k K(x_k, x) + b, \quad (11)$$

where  $\alpha_k$  are the parameters to be adjusted,  $x_k$  is the training pattern. The  $K$  function is a predefined kernel and  $x$  is the support pattern.

$$K_l = (x \bullet w), \quad (12)$$

$$K_g = e^{-\frac{\|x - \mu_j\|^2}{\sigma_j^2}}, \quad (13)$$

where  $x$  is the training pattern,  $\mu$  is the center of the Gaussian,  $\sigma$  the variance, and  $\bullet$  is the dot product.

### 3.3.3 Convolutional Neural Networks (CNNs)

CNNs are a relatively new branch of machine learning, introduced by LeCun in the 1990s [15, 16, 14]. This architecture type has turned out to be the default option for analysis, classification and treatment of image problems [29, 28, 31, 13, 11].

The success of these architectures is due to the implementation of feature extractors, trained in an automatic way. These feature extractors are convolutional filters which, by means of a stochastic gradient descent training, "learn" the characteristic features of each class. The layers of convolutional filter are called feature extraction layers. These layers also use a sampling-based discretization element, called Max-pooling, which helps to reduce the dimensionality of the elements to be classified, as well as to preserve the most predominant features.

The second element that constitutes these networks is the classification layer; usually, this layer is implemented by a Multi-Layer Perceptron (MLP) of two or three layers.

The trainable parameters that constitute this type of networks are the synaptic weights corresponding to the MLP classification layer, and the weights corresponding to the convolutional filters in the feature extraction layer. All these parameters are trained end-to-end, through optimization algorithms such as back-propagation. The general architecture of this type of network is shown in Figure 6.

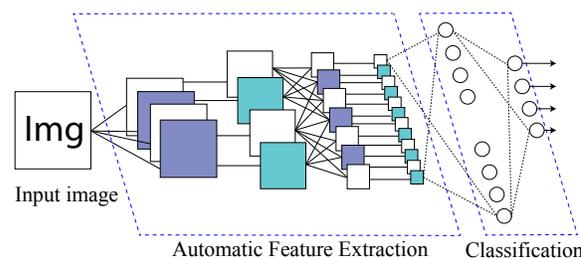


Fig. 6. Classical architecture of a CNN

## 4 Experiments and Results

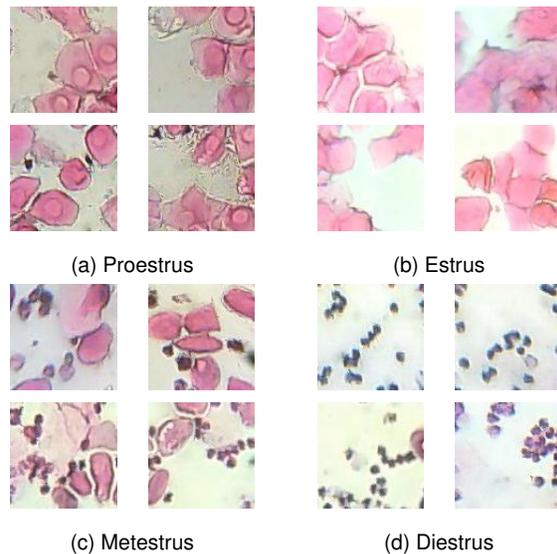
In this section, we describe the results when evaluating the different classifiers on two sets, the first composed of four stages of the estrous cycle (Figure 7) and the second by two classes. The first class is composed of the stages Proestrus and Estrus (Figures 7-a and 7-b) and the second class is formed by the stages Metestrus and Diestrus (Figures 7-c and 7-d). The second training set is shaped with the stages that define the receptivity of the female rat for its reproduction [3]. For implementation and classification details consult [10].

### 4.1 Dataset

The dataset consists of 32 images, 8 of each stage, of the estrous cycle. The images were taken with a Logitech C170 camera with a magnification of the microscope of 400x formed by an ocular 10X and an objective 40X, which helped provide a good image quality [18]. The images were divided into 16 sub-images. Each sub-image was tagged by the staff of the BUAPs vivarium Claude Bernard. The sub-images that did not present an adequate number of cells were discarded, resulting in a total of 89 images of the Diestrus stage, 125 of the Metestrus stage, 112 of the Estrus stage and 86 of the Proestrus stage with dimensions of 100x90 pixels (Figure 7).

### 4.2 Classification Methodology

The classification process is divided into two stages. In the first, we use the 105 manual features, result of texture and shape analysis



**Fig. 7.** a) The 5x5 image with four grey-levels, b) The general form of any GLCM for images with grey-levels values 0-3, c)-f) Calculation of GLCM for  $(\delta, \theta) = (1, 0), (1, 45), (1, 90), (1, 135)$

described in Section 3.2. To classify these features, we used DNN and SVM networks with radial base kernel, which are ideal for classifying this type of information.

In the second classification stage, the CNN LeNet-5 (from scratch) and pre-trained VGG16 are trained, for which the images of the cells in RGB format with a resized dimension of 150x150 are used, normalized in value, dividing the pixel value by 255.

Since we are using three different types of classifiers DNN, SVM, CNN, and to be able to have a valid comparison point between the different architectures, the original dataset was treated in the following way for all tests. From the total of 412 images, three mutually exclusive sets were generated. The first set, the training, 280 images were taken at random, 70 images for the validation set, and 62 images for the test. These last three datasets were used for the training of the DNN and SVM networks.

Due to the lack of images (412 in total and 280 for training), we chose to artificially expand the training set generating a total of 5600 images, 19

artificial images for each original image, applying the following transformations: rotation range  $[0, 180]$ , horizontal and vertical flip. At the end, the training set for the CNNs is 5600 images. In summary, only the training set for the CNNs was expanded, the validation and test sets are the same for all the trials.

#### 4.2.1 SVM Classification

For the classification of the manual features, a random search grid was generated for  $\mu$  and  $\sigma$  parameters of the radial base kernel [1]. The grid has a uniform distribution in a logarithmic scale range of  $[-10, 10]$  and a density of 10,000 samples. The classification result of this architecture is shown and compared in Section 4.3.

#### 4.2.2 DNN Classification

In the same way that a search grid of random hyper-parameters was generated for the SVM. For the DNN a similar hyper-parameter search grid was generated. This grid includes the number of intermediate layers in a range of  $[1, 6]$ , the perceptrons number per layer in a range of  $[1, 250]$ , and the learning rate in a range of  $[0.1, 0.00001]$ . The random search grid was generated with a uniform distribution of its values. The density of the grid is 10,000 elements or architectures. The classification result of this architecture is discussed in Section 4.3.

#### 4.2.3 LeNet-5 CNN Classification

Opting for a different point of view to the two previous classification methods, we decided to use a more robust architecture in terms of image classification; a Convolutional Neural Network, which contemplates both, classification and automatic feature extraction. This architecture (LeNet-5) is composed of a first block of 6 convolutional filters of 5x5, followed by a max-pooling of (2, 2) with a step of 2. The second block consists of 16 convolutional filters of 5x5, followed by a max-pooling of (2,2) with a step of 2. The third block consists of 120 convolutional filters 5x5. All the activations are ReLU functions.

The classification layer consists of a MLP of two layers, the first layer of 84 neurons and the output layer of 4 neurons for four classes classification problem, and one output perceptron neuron for a two classes classification. The results of this architecture are shown and discussed in Section 4.3.

#### 4.2.4 VGG16 Classification

In this stage and due to the results of the previous classifiers, we decided to use a pre-trained VGG16 network [29] with the Imagenet dataset [8]. This under two premises: the first is to take advantage of the previous knowledge of the pre-trained convolutional filters and, the second, using transfer learning techniques and artificial data augmentation we can combat the high percentage of overfitting presented by previous classifiers.

Two tests were carried out with the VGG16 architecture, the first using all the pretrained feature extraction blocks (convolutional and max-pooling filter layers) and training only the classifier at the end of said network. The second test was to eliminate the last feature extraction block consisting of three convolutional filters and one maxpooling.

This in order to rule out the abstraction of high-level features of the last convolutional layers, and take advantage of the low-level features of the first convolutional layers, since these features are more suitable for classifying basic elements such as cells. The MLP classifier added to this architecture consists of two layers of 100 and four neurons for four estrous cycles classification, and of 100 and one neuron for two estrous cycles classification.

### 4.3 Results

In this section, we present the results of classifying two datasets (for two and four estrous cycles), with the four architectures of neural networks as described above.

**Table 2.** Table showing the Macro and Micro F1 score for the classification of 4 estrous cycles, using 5 different classification algorithms

Network	Micro-F1	Macro-F1
MLP	0.6906	0.7059
SVM	0.7096	0.7134
LeNet-5	0.8548	0.8545
VGG-16	0.8709	0.8729
VGG-16 Modified	0.9193	0.9223

**Table 3.** Table showing the Macro and Micro F1 score for the classification of 2 sets of estrous cycles, using 5 different classification algorithms

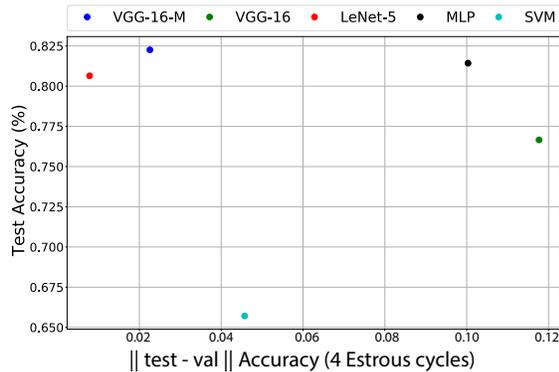
Network	Micro-F1	Macro-F1
MLP	0.8322	0.8291
SVM	0.9290	0.9281
LeNet-5	0.9193	0.9176
VGG-16	0.9516	0.9505
VGG-16 Modified	0.9838	0.9837

#### 4.3.1 Classification of 4 (Estral Cycle) Classes

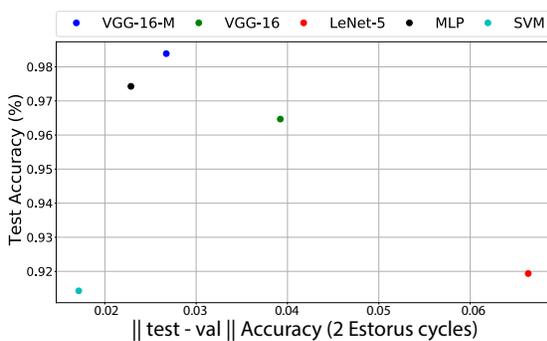
This problem turns out to be the most interesting and difficult to classify, where each class is represented by a certain estrous cycle, Proestrus, Estrus, Metestrus, and Diestrus. The results of each classifier are presented in Table 2. The percentages of classification for MLP, SVM, LeNet-5, VGG-16 and VGG-16 Modified are shown. In summary, we can conclude that for this specific problem, the convolutional neural networks have a greater ability to automatic feature abstraction, and these features are more representative than the manually selected features (see Section 3.2).

#### 4.3.2 Two-Class Classification

In this section, we present the classification results for the estrous cycles, Proestrus and Estrus stages as the first class and the second class was formed by the stages Metestrus and Diestrus.



**Fig. 8.** Graph showing the degree of overfitting of each neural network for four estrous cycles



**Fig. 9.** Graph showing the degree of overfitting of each neural network for 2 class estrous cycles

In Table 3 the Micro and Macro F1 score is shown for each neural network model. Although the classification percentages for this training set are higher than those shown in Table 2, the model with the best classification percentage again is the modified VGG-16 model.

#### 4.4 Generalization Analysis

In this section, the models with better generalization are analyzed. This is due to the fact that in the training stage and because of the training data shortages, all the aforementioned models have overfitting to a certain degree. Also, although the Micro and Macro F1 scores are reliable metrics for measuring the performance of a network, these

metrics are skewed, since they were calculated with a reduced set of only 62 images. From the above, we present a discussion on the best model to use for the classification of the estrous cycles and this is not the model that obtained the highest percentage in terms of classification.

From Figure 8, we can conclude that the architecture that presents the lower overfitting is the LeNet-5 network (shown in red). This is because it has the smallest difference between the percentages of validation and test with a classification percentage greater than 80%. Also, the use of the Modified VGG-16 network (VGG-16-M), turns out to be a good option, with a difference of 0.022% between the test and validation percentage, and a good 82% in test.

Notice that these percentages vary with respect to the percentages shown in Table 2. The latter shows the highest training rates, but they have up to 15% of overfitting, and the percentages shown in Figure 8 are the classification percentages of the architectures with the lowest degree of overfitting.

Regarding the classification of the second dataset (2 classes), from Figure 9 we can see that as far as over-training is concerned, any of the architectures used is a good option. This is because of the difference between the maximum and minimum percentage of overfitting is only 0.04 percent. This leads us to focus on the neural network with the highest percentage of classification, which in this case is the Modified VGG-16 network, which has a classification of 98.38%, and it turns out to be the best option for the dataset of 2 classes.

## 5 Conclusion and Future Work

The contribution of this work is the automatic classification of the estrous cycle through the automatic feature extraction. A methodology for the autonomous classification of the estrous cycle was presented in their four stages: Proestrus, Estrus, Metestrus, and Diestrus, as well as the classification of stages with high hormonal levels (Estrus and Proestrus) which are used for population control. In the first case, 82% of accuracy was reached while in the second case it was 98.38%.

The results obtained are considered sufficient to solve the problem of population control. However, the classifiers have problems separating the stages Metaestrous and Diestrous, so it is necessary to improve the features for these stages by increasing the number of images. In this direction, we prove that the automatic feature extraction of the CNNs is more robust than the proposed manual feature extraction in terms of generalization of neural networks, which translates into greater reliability in terms of classification. As future work, we intend to expand the dataset, in order to provide a reliable classification of the estrous cycle using convolutional neural networks as well as design an expert system based on the graphic tool proposed by Byers [3].

## Acknowledgements

E. Zamora and H. Sossa would like to acknowledge the support provided by CIC-IPN in carrying out this research. This work was economically supported by SIP- IPN (grant numbers 20180180, 20180730, 20190007), and CONACYT grant number 65 (Frontiers of Science). G. Hernández acknowledges CONACYT for the scholarship granted towards pursuing his Ph. D. studies.

## References

1. Bergstra, J. & Bengio, Y. (2012). Random search for hyper-parameter optimization. *J. Mach. Learn. Res.*, Vol. 13, No. 1, pp. 281–305.
2. Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, COLT '92, ACM, New York, NY, USA, pp. 144–152.
3. Byers, S. L., Wiles, M. V., Dunn, S. L., & Taft, R. A. (2012). Mouse estrous cycle identification tool and images. *PLOS ONE*, Vol. 7, No. 4, pp. 1–5.
4. Caligioni, C. S. (2009). Assessing reproductive status/stages in mice. *Current Protocols in Neuroscience*, Vol. 48, No. 1, pp. A.41.1–A.41.8.
5. Clausi, D. A. (2002). An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing*, Vol. 28, No. 1, pp. 45–62.
6. Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, Vol. 20, No. 3, pp. 273–297.
7. Cossio, R., Carreira, M. B., Vásquez, C. E., & Britton, G. B. (2016). Sex differences and estrous cycle effects on foreground contextual fear conditioning. *Physiology & Behavior*, Vol. 163, pp. 305–311.
8. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. *CVPR09*.
9. Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-3, No. 6, pp. 610–621.
10. Hernández, G. (2018). Estrous cycle classification through automatic feature extraction. [https://github.com/GerardoHH/Estrous\\_Cycle](https://github.com/GerardoHH/Estrous_Cycle).
11. Iandola, F., Moskewicz, M., Karayev, S., Girshick, R., Darrell, T., & Keutzer, K. (2014). Densenet: Implementing efficient convnet descriptor pyramids. *arXiv preprint arXiv:1404.1869*.
12. Kashyap, D., Somani, A., Shekhar, J., Bhan, A., Dutta, M. K., Burget, R., & Riha, K. (2016). Cervical cancer detection and classification using independent level sets and multi SVMs. *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, pp. 523–528.
13. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, pp. 1097–1105.
14. Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, Vol. 521, No. 7553, pp. 436–444.
15. Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, pp. 2278–2324.
16. LeCun, Y., Bottou, L., Orr, G. B., & Müller, K.-R. (1998). Efficient backprop. *Neural Networks: Tricks of the Trade, This Book is an Outgrowth of a 1996 NIPS Workshop*, Springer-Verlag, London, UK, UK, pp. 9–50.
17. Lee, S. Y., Lee, J.-Y., Oh, S. J., Kim, H.-C., & Kim, S. K. (2012). Expression of hepatic and ovarian antioxidant enzymes during estrous cycle in rats. *Toxicology Letters*, Vol. 212, No. 3, pp. 329–336.

18. **Marcondes, F., Bianchi, F., & Tanno, A. (2002).** Determination of the estrous cycle phases of rats: Some helpful considerations. Vol. 62, pp. 609–14.
19. **Nair, V. & Hinton, G. E. (2010).** Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML'10*, Omnipress, USA, pp. 807–814.
20. **Otsu, N. (1979).** A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 9, No. 1, pp. 62–66.
21. **Plissiti, M., Nikou, C., & Charchanti, A. (2010).** Automated detection of cell nuclei in Pap smear images using morphological reconstruction and clustering. Vol. 15, pp. 233–41.
22. **Plissiti, M. E., Nikou, C., & Charchanti, A. (2010).** Watershed-based segmentation of cell nuclei boundaries in Pap smear images. *Proceedings of the 10th IEEE International Conference on Information Technology and Applications in Biomedicine*, pp. 1–4.
23. **Plissiti, M. E., Vrigkas, M., & Nikou, C. (2015).** Segmentation of cell clusters in Pap smear images using intensity variation between superpixels. *2015 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 184–187.
24. **Priddy, B. M., Carmack, S. A., Thomas, L. C., Vendruscolo, J. C., Koob, G. F., & Vendruscolo, L. F. (2017).** Sex, strain, and estrous cycle influences on alcohol drinking in rats. *Pharmacology Biochemistry and Behavior*, Vol. 152, pp. 61–67. Sex differences on drugs affecting behavior.
25. **Quinlan, M. G., Duncan, A., Loiselle, C., Graffe, N., & Brake, W. G. (2010).** Latent inhibition is affected by phase of estrous cycle in female rats. *Brain and Cognition*, Vol. 74, No. 3, pp. 244–248.
26. **Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986).** Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. chapter Learning Internal Representations by Error Propagation. MIT Press, Cambridge, MA, USA, pp. 318–362.
27. **Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988).** Neurocomputing: Foundations of research. chapter Learning Representations by Back-propagating Errors. MIT Press, Cambridge, MA, USA, pp. 696–699.
28. **Simard, P. Y., Steinkraus, D., Platt, J. C., et al. (2003).** Best practices for convolutional neural networks applied to visual document analysis. *ICDAR*, volume 3, pp. 958–962.
29. **Simonyan, K. & Zisserman, A. (2014).** Very deep convolutional networks for large-scale image recognition.
30. **Van Der Malsburg, C. (1986).** Frank Rosenblatt: Principles of neurodynamics: Perceptrons and the theory of brain mechanisms, pp. 245–248.
31. **Wasserman, P. D. & Schwartz, T. J. (1988).** Neural networks. II. what are they and why is everybody so interested in them now? *IEEE Expert*, Vol. 3, No. 1, pp. 10–15.
32. **Zhang, L., Kong, H., Ting Chin, C., Wang, T., & Chen, S. (2013).** Cytoplasm segmentation on cervical cell images using graph cut-based approach. Vol. 23, pp. S1177–S1183.
33. **Zhao, M., Wu, A., Song, J., Sun, X., & Dong, N. (2016).** Automatic screening of cervical cells using block image processing. *Biomed Eng Online*, Vol. 15, pp. 14.

Article received on 03/12/2018; accepted on 12/04/2019.  
Corresponding author is Gerardo Hernández Hernández.