Convolutional Neural Network for Classifying the Stages of the Cell Cycle

Edgar F. Duque-Vazquez[1,1], Jonathan Cepeda-Negrete[2,2],
Joel E. López-Meza[3,3], Noe Saldana-Robles[4,1],
Raul E. Sanchez-Yanez[1,1*]

[1]Department of Agricultural Engineering, Universidad de Guanajuato,
Ex Hacienda El Copal km 9; carretera Irapuato-Silao; A.P. 311, Irapuato,
36500, Guanajuato, Mexico.
[2*]Department of Electronics Engineering, Universidad de Guanajuato,
Carretera Salamanca - Valle de Santiago km 3.5 + 1.8 Comunidad de
Palo Blanco, Salamanca, 36885, Guanajuato, Mexico.
[3]Centro Multidisciplinario de Estudios en Biotecnología de la facultad de
Medicina Veterinaria y Zootecnia, Universidad Michoacana de San
Nicolas de Hidalgo, Km 9.5 Carretera Morelia-Zinapécuaro, Posta
Veterinaria, Morelia, 58893, Michoacan, México.

*Corresponding author(s). E-mail(s): sanchezy@ugto.mx;
Contributing authors: ef.duquevazquez@ugto.mx; j.cepeda@ugto.mx;
elmeza@umich.mx; saldanar@ugto.mx;

Abstract

The cell cycle is a highly coordinated process, and its improper regulation can lead to diseases such as cancer. The detailed observation of the cell cycle constitutes a fundamental starting point for the diagnosis and prevention of diseases. One way to observe the cell cycle is through computational techniques such as deep learning, which have consolidated as tools contributing to a more precise and effective observation. Nevertheless, existing tools still exhibit a tendency for errors. In this context, the present study aims at methodological improvement through the introduction of a convolutional neural network model designed to achieve more accurate classification of cell cycle states through images. A convolutional neural network model was trained using images of Jurkat cells corresponding to different phases of the cell cycle. Subsequently, it underwent a comprehensive comparison with models designed by other researchers who also addressed the classification of cell cycle stages using the same dataset. The results

1

revealed that the proposed model demonstrated a notable ability to classify the various phases of the cell cycle. With a weighted average of 93.72% using the F1 metric, it significantly outperformed models previously documented by other authors in similar research studies. The developed model demonstrated greater consistency with the inherent characteristics of the data used, resulting in a more accurate classification of cell cycle stages. This outcome underscores the distinctive ability of a convolutional neural network to identify patterns in the cell cycle more precisely than human perception, which, at times, may be susceptible to errors.

1 Introduction

The cell cycle is a highly coordinated process that encompasses two general states: interphase and mitosis. Each of these states is divided into specific phases. The interphase is subdivided into G1, S, and G2, in that order. Mitosis, on the other hand, is divided in prophase, metaphase, anaphase, and telophase [1].

In the study of a cancer cell, the analysis of the cell cycle is of great importance. When the cell cycle is not properly regulated, a cell can transform into a cancerous state, triggering anomalies in its functioning [2]. The observation and detailed analysis of each phase of the cell cycle are essential for preventing and diagnosing these anomalies.

The study of cancer cells is commonly conducted through cell lines, which are immortalized cells designed to enable their continuous growth [3]. The Jurkat cell line, originating from acute lymphoblastic leukemia and obtained in 1977 from a child's blood, has been particularly prominent in cell cycle research.

The alterations in cellular morphology during the various phases of the cell cycle tend to be ambiguous from a visual perspective. In this context, we provide a brief explanation of the events that occur in each phase, as well as the magnitude of the visible transformations in cellular morphology in a subsequent moment.

The morphology of a cancerous cell displays variability with respect to its state in the cell cycle. During the initial phase of the cell cycle, known as G1, the cell undergoes protein and ribonucleic acid (RNA) synthesis. In this period, the cell maintains a constant morphology. The S phase constitutes the second stage of the cell cycle. At this point, the cell initiates morphological changes, exhibiting irregular shapes and undergoing an increase in size. During the third phase, known as G2, the cell initiates the duplication process of its organelles, leading to significant morphological changes compared to the two preceding phases of the cell cycle. In the prophase and metaphase stages, which belong to mitosis, no visible morphological differences are observed at a glance compared to the G2 phase. In the anaphase, notable differences become evident as the cell initiates the process of separation into two daughter cells. In the telophase, preceding cytokinesis, the cell is practically divided into two new cells, resulting in a visually evident morphological change [4].

The observation of cell alterations is carried out through microscopy techniques. The most commonly used methods in cell cycle analysis range from the use of flow cytometry [5] to the calculation of DNA content in the cell [6]. Biomarkers, such as FUCCI technology [7] are also utilized. However, it is important to note that the mentioned techniques do not provide extensive details about cell morphology in each phase of the cell cycle [8].

Currently, data analysis plays a pivotal role in virtually all areas of research. In the study of the cell cycle, there is no exception, as the application of computer vision techniques or machine learning enables the classification of cell cycle phases through images. The use of deep learning networks is the most commonly employed tool when it comes to cell cycle phase classification.

Abin et al. [9], utilized a Recurrent Neural Network in combination with Convolutional Gated Recurrent Unit layers in two of their works for cell cycle classification [9]. In both studies, a comparison was made with a ResNet model [10].

Narotamo et al. [7] utilized FUCCI technology as an indicator to determine the phase of the cell cycle. In their study, they emphasized the use of the compound DAPI for identifying the cellular nucleus. Subsequently, they employed a Support Vector Machine (SVM) for the classification of cell cycle phases. In another study of theirs [11], they proposed three approaches based on the Fast YOLO algorithm and the use of the DAPI compound for cell staining, primarily aimed at classifying phases corresponding to the interphase of the cell cycle. On the other hand, Rappez et al. [12] developed a deep neural network called DeepCycle to reconstruct the trajectory of the cell cycle. In this approach, similar to previous works, they utilized FUCCI technology.

The methodology proposed in this work is situated within a broader context and is closely related to various previous studies, especially concerning the employed dataset. In order to emphasize the current contributions in the field, the following works are presented in chronological order.

Blasi et al. [13] were the creators and the first to use the dataset in a classification. They obtained images of individual cells using flow cytometry with the aim of predicting DNA content and quantifying cell cycle phases. The method they developed proved effective for cell cycle analysis.

Eulenberg et al. [14] conducted the reconstruction of the cell cycle with the purpose of elucidating its behavior in the progression of diabetic retinopathy. For this, they implemented a neural network called DeepFlow based on MXNet, through which they classified the different states of the cell cycle, achieving an accuracy of 79.40% considering the seven states of the cycle. However, in some phases, the algorithm's classification proved to be deficient due to the limited number of available images [14].

Jin et al. [8], introduced the use of the WGAN-GP data augmentation technique to mitigate the issue of data imbalance, particularly in the phases of cell division. Subsequently, they implemented a ResNet-41 model. Considering the classification of the seven phases, they reported an accuracy of 82.10%.

Rana et al. presented two studies in which they classified the cell cycle [15][16]. Similar to Jin et al., they addressed the issue of data imbalance by employing a combination of two data augmentation techniques: WGAN and mixup. In their results, they

achieved an accuracy of 85%. In their second study, they combined three data augmentation techniques: WGAN, mixup and nonlinear mixup, resulting in an accuracy of 85.6%.

The various methodologies employed in this dataset have yielded promising results for cell cycle classification. However, a significant gap persists, presenting an opportunity for further investigation to find a methodology that classifies the cell cycle more effectively. In this study, we introduce a methodology that utilizes a convolutional neural network model capable of classifying cell cycle phases with higher accuracy than the algorithms mentioned earlier. The improvement is particularly focused on phases involving mitosis.

The document is organized into four sections following this introduction. The next section will discuss the employed methodology. Subsequently, results will be presented through images and graphs, accompanied by their discussion. Finally, the conclusions reached at the end of this study will be outlined.

2 Methodology

In this section, the methodology employed in creating the proposed Convolutional Neural Network (CNN from now on) model in this study is described. Initially, details of the dataset used are provided. Subsequently, the architecture of the developed CNN model is comprehensively outlined. Finally, a concise explanation of the methods used for comparisons is provided, accompanied by a description of the functioning of the metrics used to evaluate the performance of the developed model.

The general procedure for the development of the proposed model is illustrated in Figure 1. The procedure followed to evaluate the proposed model based on the methodology of two recent works is also depicted.

2.1 Dataset

The dataset used was extracted from the work of Blasi et al. [13]. This dataset comprises four channels (brightfield, darkfield, and fluorescence channels), each with 32,266 images. These images correspond to cancer cells from the Jurkat cell line and were captured through imaging flow cytometry.

In our study, the brightfield and darkfield channels were used. The dataset is divided into seven classes (states), corresponding to the phases of the cell cycle: G1, S, G2 (interphase), and Profase, Metafase, Anafase, and Telofase (mitosis). The distribution of images per class is shown in Figure 2. The original size of the images is 66×66 pixels; however, it has been modified to 64×64 pixels. All images are in color format.

2.2 Developed CNN model

A CNN model has been constructed following the standard conventions of the field. Essentially, a CNN is a type of artificial neural network that employs the convolution operation to extract significant features from an image.
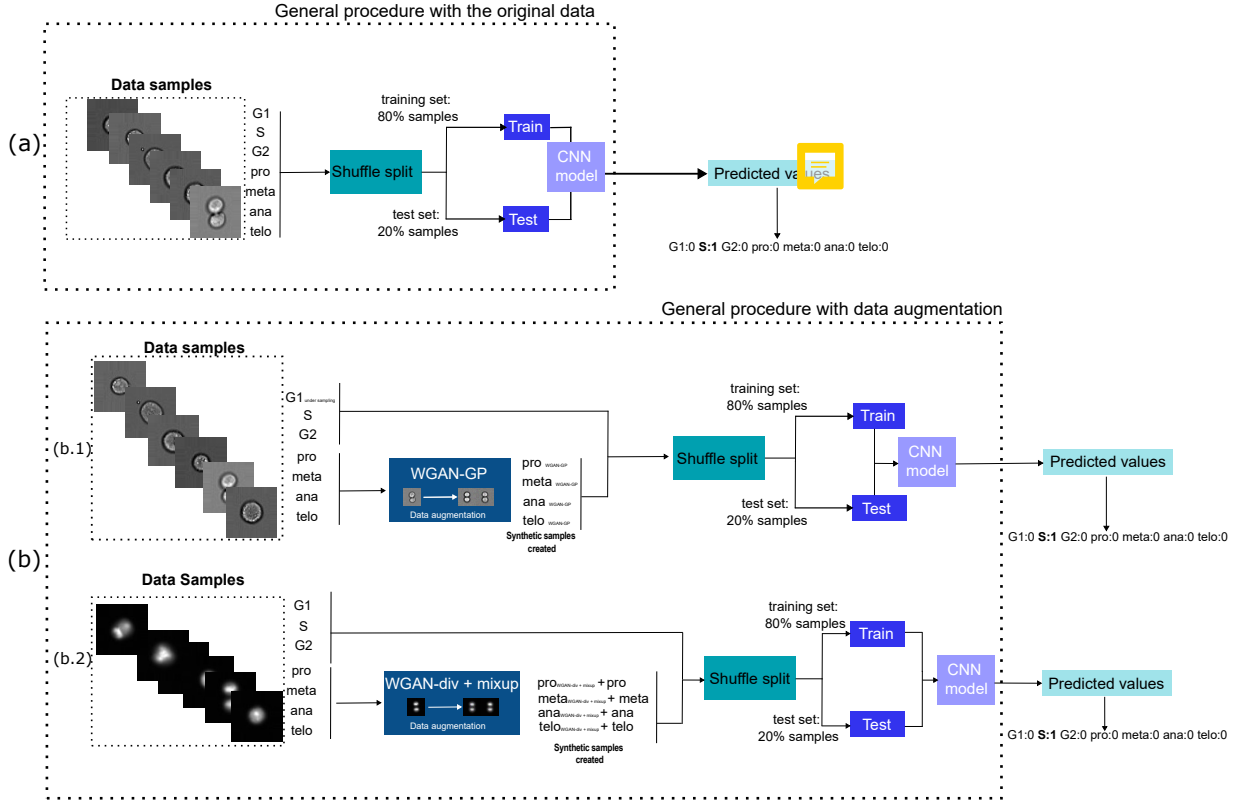
Fig. 1: General methodology used in the development of the model and its evaluation with other methodologies. (a) Methodology carried out for the developed model (b) General procedures with data augmentation (b.1) Procedure carried out with the WGAN-GP technique (b.2) Procedure carried out with the WGAN-div and mixup technique.

The developed CNN model was designed to consist of two convolutional layers, each followed by a max-pooling layer. The convolutional layers are configured with kernels of 64 and 32, which learn the features of the image. The kernel sizes are 10 and 5, respectively. The activation function used in each of these layers is the Rectified Linear Unit (ReLU). Subsequently, the flattening technique was implemented and connected to a dense layer composed of 32 neurons, followed by a dropout layer to mitigate the risk of overfitting. Then, another dense layer with 16 neurons was used; in both dense layers, the ReLU activation function was applied. Finally, a dense layer with 7 neurons was incorporated, each representing one of the cell cycle states to be classified. Consequently, a softmax activation function was applied to the last layer, facilitating the classification of the different categories (see Figure 3).
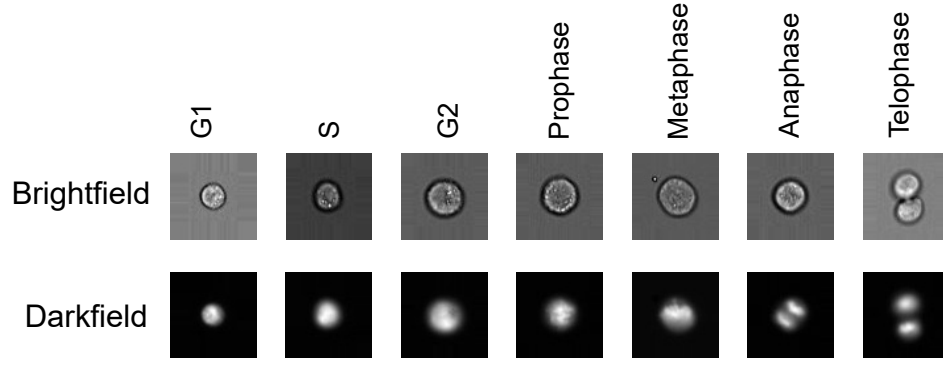
Fig. 2: An example of the cell cycle images used in this study. The images are organized by phases and categorized according to the type of imaging flow cytometry used.
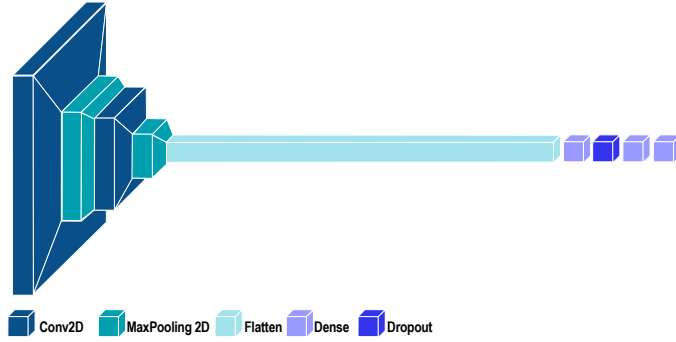
Fig. 3: Architecture of developed convolutional neural network.

2.2.1 Training model

The network was trained with the aforementioned dataset, which was split into two sets: training data (80% of the images) and test data (20% of the images). The categorical cross-entropy loss function was used, commonly employed in multiclass classification problems. Additionally, the Adam optimization function was utilized with a learning rate of 0.0001, $beta_1$=0.9, and $beta_2$=0.999, along with a batch size of 32. Finally, accuracy was used as a metric to evaluate the model's performance during both training and evaluation.

2.3 Procedure for observing the behavior of the developed CNN model

Since there were already other methodologies classifying the dataset, we decided to adopt their approaches with the aim of conducting a comparison with our results. This process allowed us to observe the behavior of the developed model. Based on the methodology proposed by Jin et al. [8], we employed the data augmentation technique

WGAN-GP [17] (see Figure 1b). This network is an improvement of a Generative Adversarial Network (GAN) that focuses on gradient penalty to address the issue of gradient fading. It stands out for its fast convergence rate and greater stability compared to its predecessor, WGAN. The WGAN-GP was trained using a batch size of 4 for Anaphase and Telophase, and 16 for Metaphase and Prophase. Training was carried out for 7000 epochs with a learning rate of 0.00001, $beta_1 = 0.01$, and $beta_2 = 0.999$. The image size was adjusted from $66 \times 66 \times 3$ to $64 \times 64 \times 3$. Network hyperparameters were maintained at their default values. Additionally, the G1 class had an excess of samples, so subsampling was applied by randomly removing images (see Figure 1). All these steps were taken to achieve a balance between mitosis and interphase class samples.

Once the new dataset was prepared, we proceeded to train the model described in Section 2.2 using the hyperparameters initially assigned. This process led to the acquisition of new performance values for the classification of cell cycle states.

On the other hand, following the methodology proposed by Rana et al. [15], we implemented the data augmentation technique WGAN-div in combination with the mixup technique see Figure 1c). Similar to WGAN-GP, WGAN-div [18] is an improvement over WGAN that approximates Wasserstein divergence to provide greater stability during training. On the other hand, mixup [19] uses the blending of two images to create a hybrid of them. For more details about these techniques, it is recommended to refer to the mentioned references. The hyperparameters used to train the WGAN-div are described below. The batch size for all classes was set to 64, with a learning rate of 0.0001 for Prophase, Metaphase, and Anaphase, and 0.00001 for Telophase. $beta_1$ was 0.01, and $beta_2$ was 0.999. Similar to WGAN-GP, the image size was adjusted from $66 \times 66 \times 3$ to $64 \times 64 \times 3$. The number of training steps per iteration for the discriminator was 10 for Prophase and 4 for Metaphase, Anaphase, and Telophase. Network hyperparameters were kept at their default values. Half of the images obtained by WGAN-GP underwent the mixup technique and were combined with images obtained by WGAN and the original images. With this new dataset, the model developed in Section 2.2 of this chapter was trained using the initially assigned hyperparameters

2.4 F1-Score

With the aim of evaluating the results of the proposed CNN model, the F1 metric was employed, which is widely recognized in the assessment of classification models. The F1-score metric combines two key metrics: precision and recall. Standard practice involves using these metrics together to provide a comprehensive assessment of model performance. The F1 score reaches its optimal value at 1 and its minimum at 0 [20]. The F1 score is calculated using Formula 1.

$$F1 = \frac{2 \cdot (Precision \cdot Recall)}{(Precision + Recall)} \tag{1}$$

where F1 is given by 2 and 3

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} \tag{2}$$

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives} \tag{3}$$

The purpose of using this metric is to obtain a measurement of the values predicted by our model compared to the labels or ground truth of each cell cycle image. Additionally, it provides a measure to compare the results of our model with those obtained through methodologies mentioned by other authors previously.

3 Results

The presented work introduces a convolutional neural network (CNN) model designed to classify the different phases of a cancerous cell cycle. This model was trained using images from the Jurkat cell line. To validate the efficiency of the developed model, a comprehensive comparison was conducted between the obtained results and those of other deep learning models utilizing the same database, as detailed in the introduction and methodology of the work. The selected metrics for model evaluation include precision, recall, the F1 score, and the weighted average of the F1 score. It is crucial to note that, during the evaluation of the proposed model, a 5-fold stratified cross-validation was employed. In this section, the evaluation and discussion are divided into two parts: the first regarding the dataset without modifications, the second concerning the data set with data augmentation using the WGAN-GP and WGAN-div + mixup techniques.

3.1 Evaluation of CNN on original data

In this study, the adoption of a Convolutional Neural Network (CNN) was based on several considerations. The utilization of a CNN is characterized by reduced complexity and heightened suitability, primarily owing to the available volume of data, which aligns seamlessly with the requisites of a moderately deep network. Furthermore, the deployment of a CNN is associated with reduced computational overhead and faster data processing.

The magnitude and heterogeneity of the training dataset constitute pivotal determinants for the efficacy of the developed model. In particular, for classes G1, S and G2, a substantial volume of data facilitates the assimilation of common patterns by the model. However, the remarkable resemblance observed among images from distinct classes presents a challenge to the model's classification accuracy, particularly concerning class S. This intricacy is discernible in the confusion matrix illustrated in Figure 4, in which the model exhibits a propensity to misclassify class S as G1 and G2, thereby contributing to a reduction in overall model accuracy. Conversely, the classes associated with mitosis exhibit a significant scarcity of images in both the training and validation sets. This data limitation adds an additional level of complexity to the model's task, as it fails to adequately grasp the specific patterns inherent in these categories.

Due to inherent characteristics in the employed data, such as the disparity in the number of samples per class and the overall dataset size, the neural network tends to experience overfitting in those classes with higher representation, leading
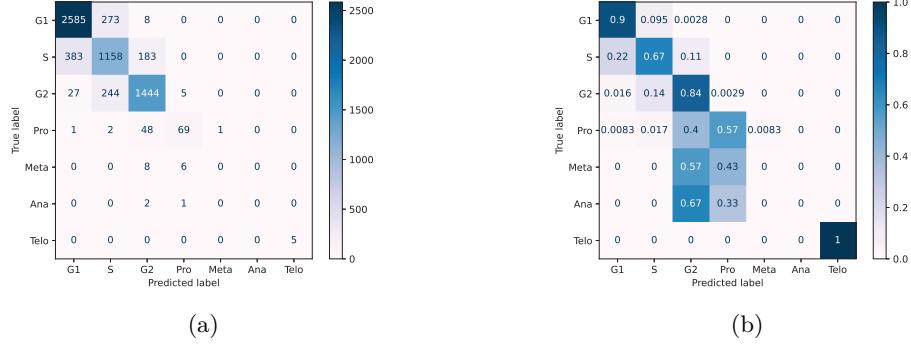
Fig. 4: (a) Confusion matrix of the classification of the original dataset. (b) Normalized confusion matrix of the classification of the original dataset.

to poor performance and directly impacting the overall model performance. Model training was halted at epoch 22, considering the observed variations in the loss function on the validation data. An increase in the loss function of the validation data is evident, contrary to expectations, and a decrease in accuracy is observed instead of an improvement.

In order to assess the performance of the proposed model, a detailed comparison was conducted, as illustrated in Table 1, between the outcomes achieved by our model and those obtained by Jin et al [8]. In this table, it becomes evident that the imbalance in the distribution of data per class emerges as a determining factor in the model's performance. Positive outcomes stand out in the G1, G2, and Telophase stages, largely attributable to the abundance of data and the clarity of the object to be classified in the images, in contrast to other phases of the cell cycle. The metrics reveal a significant difference favoring our model compared to that developed by Jin et al. [8]. This analysis highlights the crucial importance of managing the imbalance in the quantity of data per class to enhance the model's generalization. The quality and quantity of available data are fundamental elements that directly impact the model's ability to effectively classify the various phases of the cell cycle.

Table 1: Comparison of the classification results between our model and those of Jin et al., using the original dataset.

Phase	Precision		Recall		F1	
	Jin et al.'s model	Proposed model	Jin et al.'s model	Proposed model	Jin et al.'s model	Proposed model
G1	0.8316	**0.8628**	0.8403	**0.9020**	0.8359	**0.8820**
S	0.6765	**0.6905**	**0.7052**	0.6717	**0.6905**	0.6810
G2	0.8453	**0.8529**	0.8012	**0.8395**	0.8241	**0.8462**
Prophase	**0.8521**	0.8519	**1.0000**	0.5702	**0.9202**	0.6832
Metaphase	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Anaphase	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Telophase	0.0000	**1.0000**	0.0000	**1.0000**	0.0000	**1.0000**
Weighted Average	0.7835	**0.8117**	0.7844	**0.8152**	0.7853	**0.8127**

9

3.2 Results of the CNN architecture trained with data augmentation

To address the data imbalance, Jin et al., proposed the implementation of the WGAN-GP data augmentation technique, coupled with undersampling in the G1 phase. The methodology of by Jin et al. was adopted to assess the performance of the proposed CNN model, now incorporating data augmentation techniques.

Figure 5 displays the depicting results in a confusion matrix of the model trained with data augmentation and G1 undersampling. Recalling the challenges encountered with imbalanced data, an increase in the number of true positives in class S is observed, while the count of true negatives and false negatives has slightly decreased. Furthermore, in states associated with mitosis, a notable increase in true positives is evident, a consequence of incorporating new data provided by the data augmentation technique.

           (a)   (b)
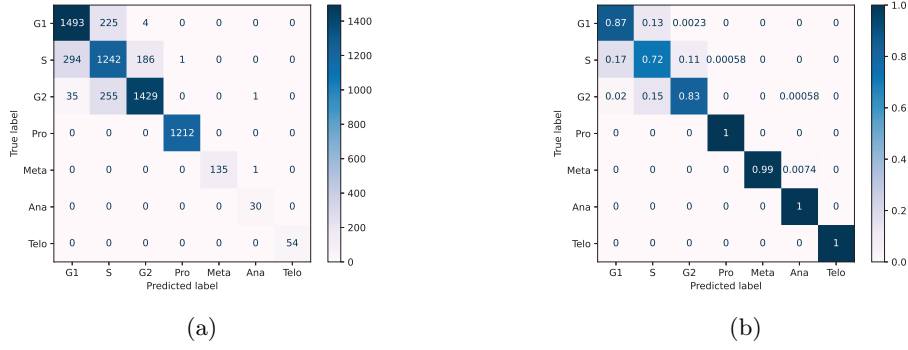
Fig. 5: (a) Confusion matrix of the classification of the dataset using the WGAN-GP data augmentation. (b) Normalized confusion matrix of the classification of the dataset using the WGAN-GP data augmentation.

Table 2 showcases improvements in most phases in terms of metrics, unlike the imbalanced dataset, for both the methodology proposed by Jin et al. and the proposed CNN model. The results of our model, in particular, are superior, suggesting that the utilization of a conventional CNN model can yield classification results similar to or even better than other deep learning architectures.

Rana et al. conducted a study in which they trained a ResNet using images generated through the combination of data augmentation techniques WGAN-div + mixup. To assess the performance of our CNN model in this work, we relied upon the methodology proposed in [15]. The proposed CNN was trained using data generated through the mentioned augmentation techniques, achieving superior performance. This contrasts with the approach taken in [8], highlighting the model's strong performance both on data without the application of augmentation techniques and on data with augmentation.

Table 2: Class-wise comparison between the proposed model and Jin et al. using the dataset with WGAN-GP data augmentation and oversampling in the G1 stage.

Phase	Precision		Recall		F1	
	Jin et al.'s model	Proposed model	Jin et al.'s model	Proposed model	Jin et al.'s model	Proposed model
G1	0.8181	**0.8194**	0.8490	**0.8670**	0.8333	**0.8426**
S	0.6700	**0.7213**	0.6918	**0.7208**	0.6808	**0.7210**
G2	0.8456	**0.8826**	0.7895	**0.8308**	0.8166	**0.8559**
Prophase	0.9934	**0.9992**	0.9909	**1.0000**	0.9922	**0.9996**
Metaphase	0.8125	**1.0000**	0.9559	**0.9926**	0.8784	**0.9963**
Anaphase	**1.0000**	0.9375	0.0667	**1.0000**	0.1250	**0.9677**
Telophase	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Weighted Average	0.8210	**0.8490**	0.8184	**0.8481**	0.8174	**0.8481**

In the confusion matrix of the Figure 6, a notable scarcity of false positives and false negatives is evident. Specifically focusing on classes associated with mitosis, a virtually flawless classification is achieved, thanks to the implementation of data augmentation techniques. It is crucial to consider that, superficially, mitotic states may exhibit variations, underscoring the importance of differentiation through the inclusion of more examples.

(a)

(b)
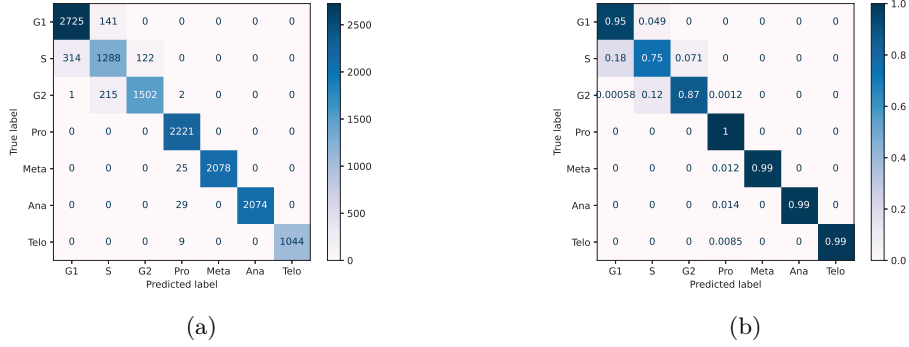
Fig. 6: (a) Confusion matrix of the classification of the dataset using the WGAN-GP data augmentation. (b) Normalized confusion matrix of the classification of the dataset using the WGAN-GP data augmentation.

In general, we can affirm that the developed model, fueled by examples provided by a WGAN-div, demonstrated superior performance. In Table 3, a comparison with the methodology proposed in [15]. is observed, focusing specifically on the F1 metric, where the algorithm shows good results in the majority of cell cycle states. With this, we can confirm that the proposed model enhances its efficiency by incorporating additional data.

The results show significant improvements in each of the classes. As anticipated, the G1, S, and G2 phases exhibit notable similarities. However, the most significant differences are observed in the classification of phases associated with mitosis, with metaphase being the most affected. This is attributed to the achieved differentiation

Table 3: Class-wise comparison between the proposed model and Rana et al.'s [15] using the dataset with WGAN-div data augmentation and mixup.

Phase	Rana et al.'s model	Proposed model
G1	**0.9300**	0.9280
S	0.7500	**0.7648**
G2	0.8400	**0.8983**
Prophase	0.9200	**0.9856**
Metaphase	0.5700	**0.9940**
Anaphase	**1.0000**	0.9931
Telophase	**1.0000**	0.9951
Weighted Average	0.8600	**0.9372**

between prophase and telophase classes, which visually resemble metaphase the most. The weighted average also reflects a higher number of successful predictions per phase in favor of the proposed model, indicating superior overall performance

4 Conclusions

This project successfully achieved the classification of the 7 states in the cell cycle of a cancerous cell. It was demonstrated that a standard CNN model is capable of effectively classifying cell cycle phases even in an imbalanced dataset. The results indicate that the use of a conventional CNN can yield similar or even better classification outcomes without the need for an excessively deep network structure. It is also noteworthy that the efficiency of the CNN model improves when data augmentation is applied. Cell cycle phases, which can be challenging for the human eye, can be accurately classified using a deep learning model. Additionally, data augmentation techniques prove invaluable when a dataset lacks a sufficient number of images per class. This approach allows for more robust model training, resulting in enhanced performance when evaluating the model's accuracy with diverse data.

Declarations

References

[1] Caglar, H.O., Biray Avci, C.: Alterations of cell cycle genes in cancer: unmasking the role of cancer stem cells. Molecular biology reports **47**, 3065–3076 (2020)

[2] Coffman, J.A.: Cell cycle development. Developmental cell **6**(3), 321–327 (2004)

[3] Drexler, H.G., et al.: The leukemia-lymphoma cell line factsbook. Technical report, Academic Press (2000)

[4] Needham, D.: Possible role of cell cycle-dependent morphology, geometry, and mechanical properties in tumor cell metastasis. Cell biophysics **18**, 99–121 (1991)

[5] FANG, H.-S., LANG, M.-F., SUN, J.: New methods for cell cycle analysis. Chinese Journal of Analytical Chemistry **47**(9), 1293–1301 (2019) https://doi.org/10.1016/S1872-2040(19)61186-2

[6] Roukos, V., Pegoraro, G., Voss, T.C., Misteli, T.: Cell cycle staging of individual cells by fluorescence microscopy. Nature protocols **10**(2), 334–348 (2015)

[7] Narotamo, H., Fernandes, M.S., Moreira, A.M., Melo, S., Seruca, R., Silveira, M., Sanches, J.M.: A machine learning approach for single cell interphase cell cycle staging. Scientific Reports **11**(1), 19278 (2021)

[8] Jin, X., Zou, Y., Huang, Z.: An imbalanced image classification method for the cell cycle phase. Information **12**(6), 249 (2021)

[9] Jose, A., Roy, R..i.-.e.j., Stegmaier, J.: Weakly-supervised temporal segmentation of cell-cycle stages with center-cell focus using recurrent neural networks. In: BVM Workshop, pp. 212–219 (2023). Springer

[10] Jose, A., Roy, R., Eschweiler, D., Laube, I., Azad, R., Moreno-Andrés, D., Stegmaier, J.: End-to-end classification of cell-cycle stages with center-cell focus tracker using recurrent neural networks. In: ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–5 (2023). IEEE

[11] Narotamo, H., Fernandes, M.S., Sanches, J.M., Silveira, M.: Interphase cell cycle staging using deep learning. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 1432–1435 (2020). IEEE

[12] Rappez, L., Rakhlin, A., Rigopoulos, A., Nikolenko, S., Alexandrov, T.: Deepcycle reconstructs a cyclic cell cycle trajectory from unsegmented cell images using convolutional neural networks. Molecular systems biology **16**(10), 9474 (2020)

[13] Blasi, T., Hennig, H., Summers, H.D., Theis, F.J., Cerveira, J., Patterson, J.O., Davies, D., Filby, A., Carpenter, A.E., Rees, P.: Label-free cell cycle analysis for high-throughput imaging flow cytometry. Nature communications **7**(1), 10256 (2016)

[14] Eulenberg, P., Köhler, N., Blasi, T., Filby, A., Carpenter, A.E., Rees, P., Theis, F.J., Wolf, F.A.: Reconstructing cell cycle and disease progression using deep learning. Nature communications **8**(1), 463 (2017)

[15] Rana, P., Sowmya, A., Meijering, E., Song, Y.: Imbalanced cell-cycle classification using wgan-div and mixup. In: 2022 IEEE 19th International Symposium on

Biomedical Imaging (ISBI), pp. 1–4 (2022). IEEE

[16] Rana, P., Sowmya, A., Meijering, E., Song, Y.: Data augmentation for imbalanced blood cell image classification. bioRxiv, 2022–08 (2022)

[17] Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C.: Improved training of wasserstein gans. Advances in neural information processing systems **30** (2017)

[18] Wu, J., Huang, Z., Thoma, J., Acharya, D., Van Gool, L.: Wasserstein divergence for gans. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 653–668 (2018)

[19] Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412 (2017)

[20] Rijsbergen, C.: Information retrieval 2nd ed buttersworth. London, 115 (1979)