Convolutional Neural Network U-Net for Trypanosoma cruzi Segmentation

Allan Ojeda-Pat[1]([✉]) [iD], Anabel Martin-Gonzalez[1] [iD], and Roger Soberanis-Mukul[2] [iD]

[1] Computational Learning and Imaging Research, Facultad de Matematicas,
Universidad Autonoma de Yucatan, Merida, Mexico
allan.ojedaa@gmail.com, amarting@correo.uady.mx
[2] Computer Aided Medical Procedures, Technical University Munich, Munich, Germany
rodsom22@gmail.com

Abstract. Chagas disease is a mortal silent illness caused by the parasite *Trypanosoma cruzi* that affects many people worldwide. A blood test is one of the preferred methods to get an accurate diagnosis of the disease but takes a long time and requires too much effort from the experts to analyze blood samples in the search of the parasites presence. Therefore, it is very useful to have an automatic system to detect the parasite in blood sample microscopic images. In this paper we present a deep learning method to segment the *T. cruzi* parasite on blood sample images. We implemented a convolutional neural network based on the U-Net model and we trained it with different loss functions to get accurate results. We report an F2 value of 0.8013, a recall value of 0.8702, a precision value of 0.6304 and a Dice score value of 0.6825.

Keywords: Chagas disease · Deep learning · U-Net segmentation

1 Introduction

According to the World Health Organization, Chagas disease is a potentially dangerous disease caused by the parasite *Trypanosoma cruzi* [1]. This parasite resides mainly in Latin America and is transmitted to humans by the feces of triatomine insects, colloquially known as kissing bugs or insect 'pic'. It is estimated that there are about 7 to 8 million people infected with Chagas disease in the world, most of which are in Latin America [1, 2], and more than 10 thousand deaths per year are attributed [2]. It is estimated that more than 30% of infected patients suffer from heart problems, and above 10% suffer from digestive, neurological or mixed problems [1]. Many of these symptoms do not occur until years after infection with the parasite. It is important to mention that during the initial phase of Chagas disease, which occurs in the first two months after infection, many parasites circulate in the blood. After the initial phase, the chronic phase continues. In this phase the parasites lodge mainly in the heart and in the digestive muscles, making it difficult to locate them in a common blood sample through a microscope [3]. It is important to carry out the detection studies during the initial phase for effective diagnosis and prompt treatment of the disease [4]. A blood test

is essential to detect and treat Chagas disease in time. The most common tests for the diagnosis of Chagas disease include ELISA test and by blood smear inspection. A blood smear involves placing a drop of blood on a slide. The sample is stained for later analysis under the microscope. Detection of parasites through microscopic inspection of blood smears is a very common and convenient technique, but sometimes it is a process that requires a lot of time and effort when many blood samples have to be analyzed [5–9]. In recent years, Machine learning and Deep learning techniques have been rapidly used in the medical imaging research area to diagnose a disease [10], and these techniques could help to reduce the effort of getting a quick diagnosis of the Chagas disease. With an automated segmentation system, experts could easily confirm if it is a proper identification/diagnosis of the *T. cruzi* parasite and experts can get relevant data about sizes and shapes for further use, all of this with saving time. The purpose of this article is to present an automated method based on deep learning to segment the *T. cruzi* parasite on blood samples images as an alternative to a tedious common manual inspection. With this segmentation method, experts can easily confirm and verify if the parasite is present in the blood sample images and collect information from the segmented parasites.

Following in Sect. 2 we describe the state of the art of clinical and machine learning works for Chagas disease detection. Section 3 explains the image dataset and manual work. Section 4 details the method developed with the U-Net model and the details of the training. Section 5 details the results of the training and the conclusions are presented in Sect. 6.

2   Previous Work

In the state of the art for Chagas disease diagnosis, we can separate the works on two groups: the clinical methods and the Machine learning-based methods.

2.1   Clinical Methods

A clinical method requires the use of a specialized portable test for the diagnosis of the Chagas disease. One of them is called Chagas Stat-Pak presented in [11] which has 99.6% sensitivity and 99.9% specificity; this performance is comparable with the obtained with an ELISA test [3]. Another method that requires a portable test is called Chagas Detect Plus [12]. Although these portable tests for Chagas detection are quick and precise, it is mandatory to confirm with a manual or serological method to get an accurate diagnosis [12]. An automated system based on Machine learning and/or Deep learning could replace or assist the manual confirmation method, and in case of a positive result, the patient can begin the appropriate treatment.

2.2   Machine Learning Methods

In the machine learning and computer vision area, some studies have been reported for the detection of the *Trypanosoma cruzi* parasite in blood sample images [3, 4, 8, 13].

Detection Methods. In 2013, Uc-Cetina et al. [4], propose the detection of the *T. cruzi* parasite in blood samples images using an algorithm based on the analysis of the Gaussian discriminant. In the classification of a new set of input pixels, they calculate the probabilities of the Bayesian model, and the highest probability will indicate the dominant class of the sample. As the final step, they implemented a search algorithm to find possible candidates of parasites to extract and classify the input vector. The performance rates they reported are false negatives 0.0167, false positives 0.1563, true negatives 0.8437 and true positives 0.9833.

In 2013, in the work of Soberanis-Mukul et al. [8], an automated algorithm is proposed to detect Chagas disease using different segmentation and classification techniques. In the first step, a binary segmentation is applied using a Gaussian classifier. By calculating the probabilities of the Bayesian model, it is segmented groups of pixels that could represent a parasite. Later, the segmented images go through a K-Nearest-Neighbors classifier [16] previously trained to get a binary classification: it is a *T. cruzi* parasite or not. The authors reported a sensitivity value of 98% and a specificity value of 85%.

In 2015, in the work of Uc-Cetina et al. [3], it is compared two classify algorithms: AdaBoost and Support Vector Machine (SVM) for the detection of *T. cruzi* parasite on blood sample images. The proposed algorithm consists of the detection of possible parasites through training of a binary AdaBoost classifier feed up with features extracted by using specific Haar templates designed for the morphology of the *T. cruzi* parasite. A post-processing step is applied using an SVM classifier to assist in discard false positives. The authors compared the proposed AdaBoost + SVM algorithm with a single SVM algorithm. The classification using AdaBoost + SVM obtained the best results reporting a sensitivity value of 100% and a specificity value of 93.25%.

Segmentation Methods. In 2014, Soberanis-Mukul [13] made a comparison of three classifiers: SVM, AdaBoost and Artificial Neural Networks (ANN) for detection and segmentation of the *T. cruzi* parasite. His methodology consists of three steps: computing of superpixels, extraction of optimal features and the training of the classifiers (ANN with backpropagation algorithm, AdaBoost with perceptron assembly and SVM). Superpixels are formed by sets of pixels that describe continuous regions delimited throughout the image. The proposal to use groups of pixels allows features to be extracted considering the neighborhood and its relationship within a group. In the experimentation stage, the author compares his three classifiers with three classifiers of the state of the art (Gaussian and Bayes Classifier and the algorithm proposed in [14]). In the end, the author reports that the Gaussian classifier had the lowest mean square error with 0.18568, followed by the SVM classifier proposed by the author who obtained an error of 0.22635, and finally, the ANN classifier with an error of 0.361.

3 Dataset

The database consists of 974 color images on RGB format of size 2560×1920 pixels.

The images were taken from blood samples of mice infected with the *T. cruzi* parasite. The blood samples show blood cells and *T. cruzi* parasites highlighted with special ink
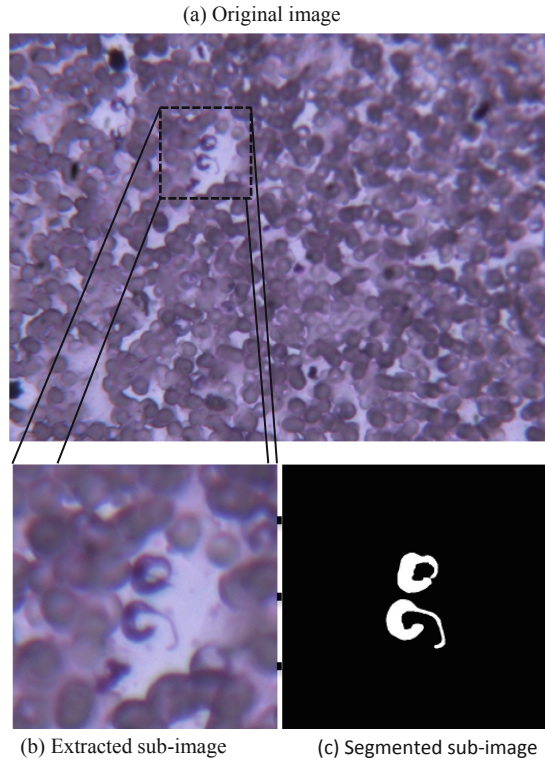
(a) Original image

(b) Extracted sub-image   (c) Segmented sub-image

Fig. 1. Ground truth generation of a sub-image. In (a) we extract (b), and (c) is the manual segmentation of (b).

for easy recognition. Some of the original images contain multiple parasites, and other images do not contain any parasites. As a first step to create our dataset, we cropped 940 sub-images of size 512×512 pixels with contain parasites presence, and 60 sub-images of size 512×512 pixels without parasites taken from the original database.

To train the segmentation network with ground truth examples, it is necessary to get a manual binary segmentation of every sub-images. An example of the manual work can be observed in Fig. 1. In the segmented sub-images, white pixels represent a parasite and black pixels represent the background.
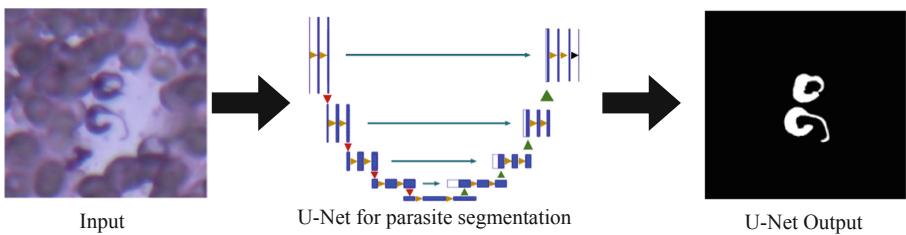

Input   U-Net for parasite segmentation   U-Net Output

Fig. 2. The proposed method for *T. cruzi* parasite segmentation

4   Method

This section introduces the method used to segment the *T. cruzi* parasite.

In computer vision, the goal of semantic segmentation is to label each pixel of an image with its corresponding class. The desired output is an image where all the pixels are classified using, for example, a convolutional neural network (CNN). We can define the above as

$$p = g(x) \tag{1}$$

where x is an RGB input image, p is a segmentation output and $g()$ is a CNN. Figure 2 shows the complete visual method.

4.1   U-Net Model

The U-Net model is a fully convolutional neural network, i.e., an architecture with only convolutional layers, to perform semantic segmentation of mainly biomedical images. The U-Net was developed by Ronneberger et al. [15]. The architecture has two paths. The first path is the contraction or the encoder path. The encoder captures the context in the image through a stack of convolutional and down-sampling layers. The second path is the symmetric expanding path or the decoder. The decoder enables precise location of the classified pixels through up-convolutions. More technically, the contraction path consists of the repeated application of two 3×3 convolutions, each one followed by a Rectified Linear Unit (ReLU) activation function, and one 2×2 max-pooling operation after two convolutional layers. The previous process is repeated five times increasing the number of convolutional filters (64, 128, 256, 512 and 1024). To get fine details in the segmentation map, multi-scale features are combined by connecting corresponding resolutions in the contracting and expanding path. The expanding path consists of the repeated application of an up-sampling operation (through one 2×2 transposed convolution) followed by a concatenation operation (with the corresponding feature map from the contracting path), and two 3×3 convolutions with a ReLu activation function. The previous process is repeated four times decreasing the number of convolutional filters (512, 256, 128 and 64), and at the end, a 1×1 convolution (with two filters) is applied to get the final segmentation map. The entire network has 23 convolutional layers. For the purpose of this work, we used a variation of the U-Net architecture by applying zero-padding to every convolution operation to maintain the input image dimensions; and a 1×1 convolution (with one filter) applied to the last layer to get the binary segmentation map, obtaining a CNN with 24 layers. Figure 3 shows the proposed U-Net model.

4.2   Training

The complete dataset consists of 1000 images, each one with its corresponding ground truth image. The training set consists of 600 images of which 564 images contain *T. cruzi* parasites and 36 images do not contain any parasite. The validation set consists of 200 images of which 188 images contain *T. cruzi* parasites and 12 images do not contain
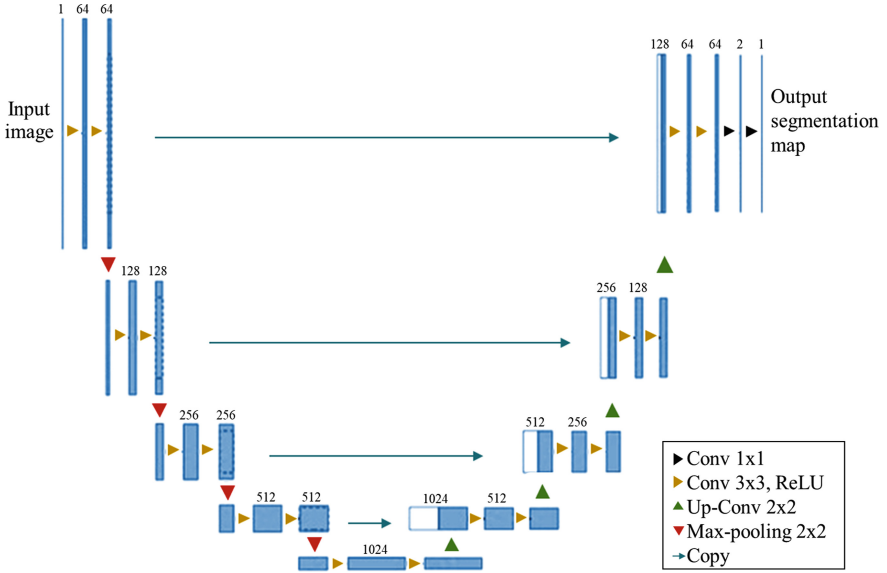
Fig. 3. U-Net model proposed with an RGB input image and a segmentation map as output.

any parasite. The final test set consists of 200 images with the same distribution as the validation set.

As shown in Table 1, the dataset has a very high-class imbalance, i.e., most of the pixels of each image correspond to the background and only a small number of pixels correspond to *T. cruzi* parasite class. During training, if there is a high-class imbalance on the dataset, the network tends to predict the most common class [9]. The output image may contain most pixels of the predominant background class.

Table 1. Pixels distribution per class of the dataset

Class	*T. cruzi* parasite	Background
Pixels	1.8%	98.2%

To mitigate this problem, we train the network with a custom loss function called Weighted Binary Cross-Entropy Loss (WBCE) [16]. This loss function is defined in terms of the Binary Cross-Entropy Loss (BCE):

$$BCE(p, \hat{p}) = -\big(p\,log(\hat{p}) + (1 - p)\log(1 - \hat{p})\big) \tag{2}$$

where $Y = 1$ is the *T. cruzi* parasite class, $Y = 0$ is the background class, $P(Y = 1) = p$ and $P(Y = 0) = 1 - p$ are the values of the ground truth for both classes. The predicted

probability for class 1 is given by the sigmoid function $P\left(\hat{Y} = 1\right) = \frac{1}{1+e^{-x}} = \hat{p}$ and for class 0 is $P\left(\hat{Y} = 0\right) = 1 - \hat{p}$.

The Weighted Binary Cross-Entropy Loss [16] is defined as:

$$WBCE(p, \hat{p}) = -\left(\beta p \log(\hat{p}) + \alpha(1 - p)\log(1 - \hat{p})\right) \tag{3}$$

where β represents the proportion of the *T. cruzi* parasite class and α the proportion of the background class.

We used data augmentation to increase the number of training samples by applying random transformations such as zooming, rotations, vertical and horizontal flips to both the input and ground truth images. Thus, we increase 3 times the training dataset size.

We conduct two experiments. The first experiment consists on training the U-Net model for 48 epochs with the Weighted Binary Cross-Entropy Loss function, steps per epoch value of 600, He initialization [17], Adam Optimizer as an optimizer algorithm with a learning rate of 1e−4, batch size of 2, $\beta = 9.0$ and $\alpha = 1.0$ (given more weight to *T. cruzi* parasite class). It takes approximately 14 h to train on the Google CoLab platform [18] using Keras framework.

The second experiment consists on training the U-Net model for 62 epochs with the regular Binary Cross-Entropy Loss, steps per epoch value of 300, He initialization, Adam Optimizer with a learning rate of 1e-4 and bach size of 2. It takes approximately 15 h to train on the same platform.

We use the Dice Coefficient value as a metric to monitor how good the predictions are compared with the ground truth. The Dice Coefficient is defined in [15] as:

$$Dice\,Coefficient = \frac{2 * |P \cap Y|}{|P| + |Y|} \tag{4}$$

where P is the segmentation map returned by the U-Net and Y is the ground truth image. The Dice Coefficient compares the pixel-wise agreement between P and Y. The metric returns a value close to one if the predictions of the network are close enough.

5   Results

The learning curves of the WBCE loss function for the first experiment is shown in Fig. 4. Figure 5 shows its mean Dice Coefficient values of every epoch during training, where on epoch 41, we get the higher Dice Coefficient value of 0.6735 for the validation set.

As we can see in Fig. 4, the model reaches a local minimum in the validation loss and it cannot decrease its validation error in contrast with the training error. Thus, we can conclude that there is a slight overfitting (Fig. 4) that does not affect the segmentation performance as we can see a higher Dice Score value on epoch 41 (Fig. 5).
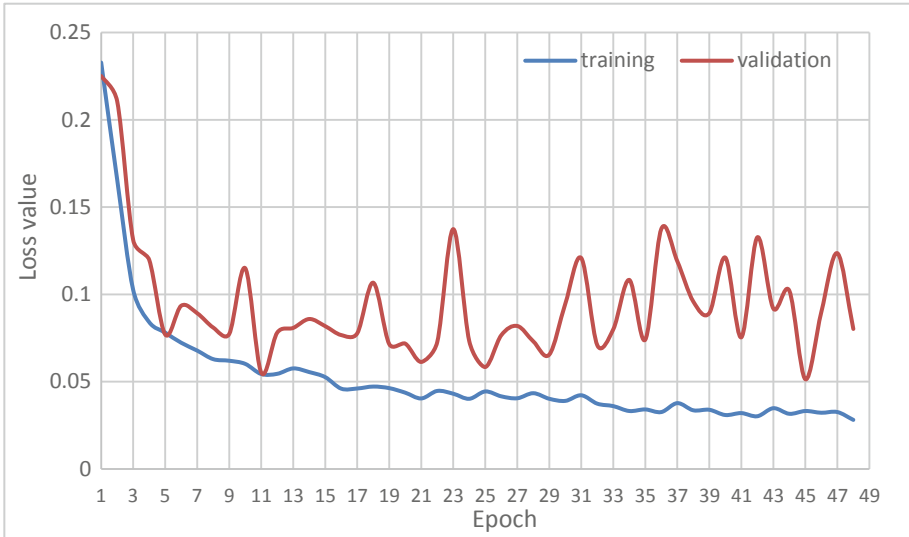
Fig. 4. Training curves of U-Net with WBCE Loss
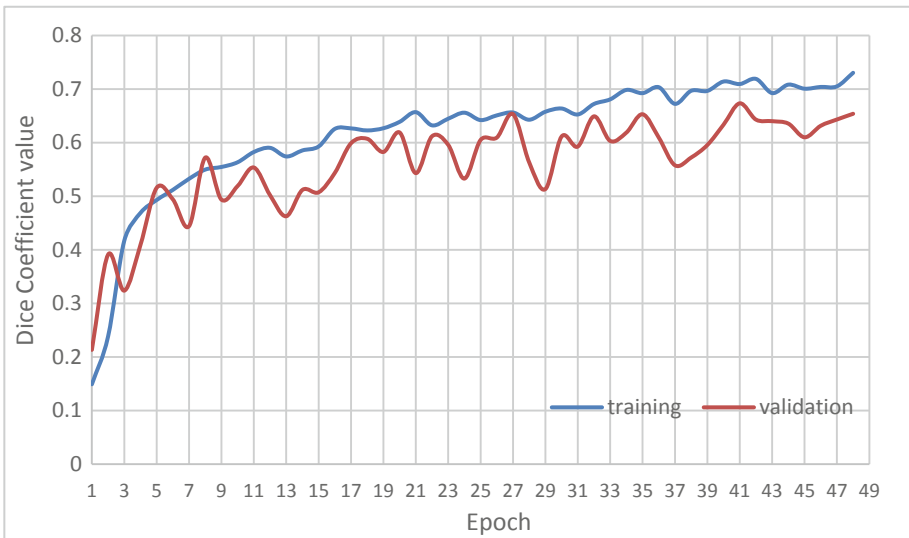

Fig. 5. Dice Coefficient values of U-Net with WBCE Loss

Figure 6 shows the learning curves of the BCE loss function for the second experiment and Fig. 7 shows its mean Dice Coefficient values.

On epoch 54, we get the higher Dice Coefficient value of 0.733 for the validation set. In Fig. 6, we can see that the model cannot decrease the validation loss after 62 epochs and tends to a horizontal asymptote, we conclude that it has to be result of the model reaching a local minimum.
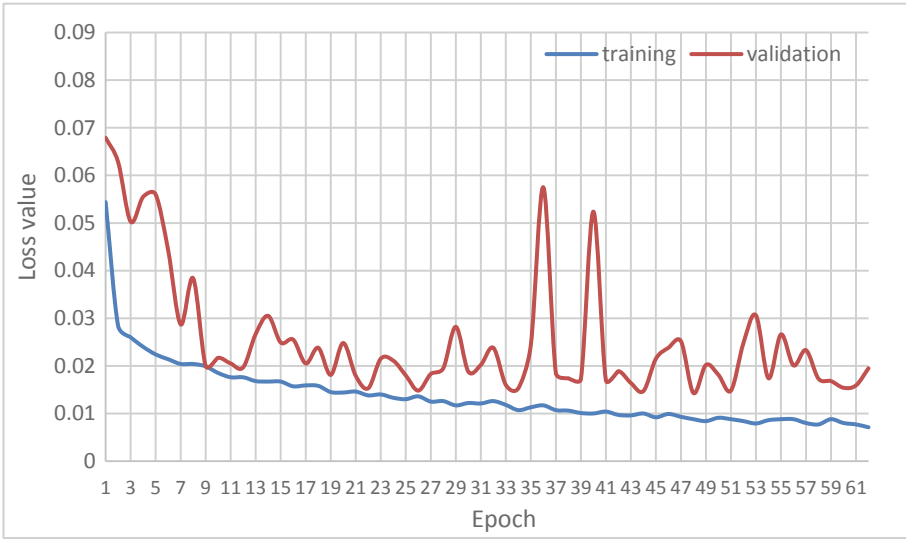
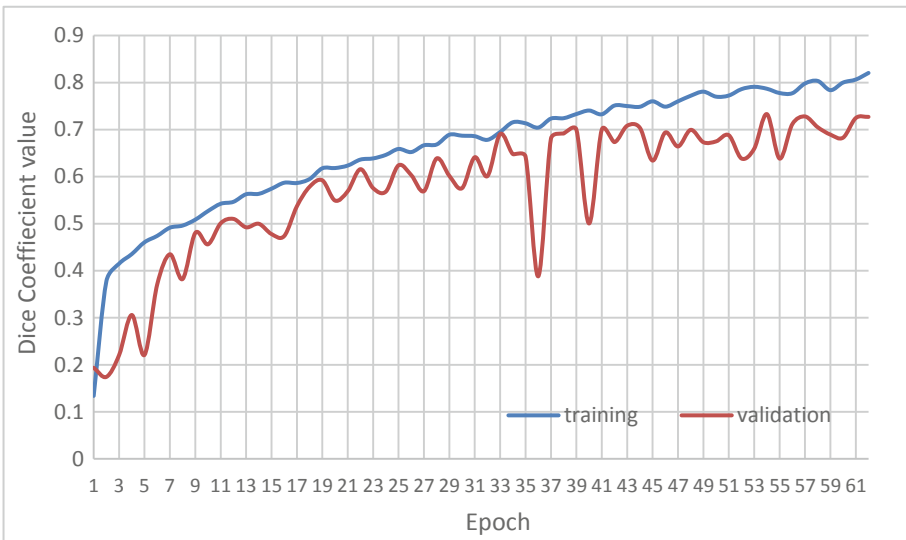**Fig. 6.** Training curves of U-Net with BCE Loss

Fig. 7. Dice Coefficient values of U-Net with BCE Loss

We compare the results obtained with the F2 score. The F2 score is a good choice as we need to measure how good the final binary segmentation is in terms of how many correct positive classifications (regions of pixels of *T. cruzi* parasite) are made. The F2 score measures the accuracy of a test in terms of the precision and recall values [19].

The F2 score is defined as:

$$F2\,score = 5 \cdot \frac{precision \cdot recall}{(4 \cdot precison) + recall} \tag{4}$$

The precision is the ratio of true positives (TP) to all predicted positives, true positives and false positives (TP + FP), as in:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

The recall is the ratio of true positives (TP) to all actual positives, true positives and false negatives (TP + FN), as in:

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

Table 2. Results obtained from test set

Model	Loss function	Loss value	Dice Coefficient value	Recall	Precision	F2 score
1	WBCE	0.0588	0.6825	0.8702	0.6304	0.8013
2	BCE	0.0148	0.7072	0.7808	0.7671	0.7714

As we are focus on the *T. cruzi* parasite segmentation, we are interested in the correct classification of class 1 pixels. Analyzing the results presented in Table 2 for the final test set, we can assume that model 1 trained with the WBCE performs better on classifying positive pixels based on the F2 score. A high recall value means that the model is focused on predicting as many correct positive pixels as possible to keep the shape of the parasite, although this could result on more false positives, as we can see in the lower precision value. Even though the Dice Coefficient value is slightly better on the model 2, it only means that the pixel-wise agreement is good to both classes of pixels (*T. cruzi* and background pixels), but its lower F2 score means that it does not predict the positive pixels as well as the model 1 does.

In Fig. 8, we can observe some of the qualitative results for both experiments. As we can see, the predictions are very closed to the ground truth, although there could be a room of improvement for tricky images.

In Fig. 9, we can see five examples of wrong predictions or not-complete segmentations returned by the U-Net. Analyzing these resulting images, we conclude that shape orientation and morphology of the *T. cruzi* parasites presented in Fig. 9c to e are not the most common on the training set, and as a result the U-Net does not learn how to properly segment them. In Fig. 9b, the reason for a black segmentation output is because the input image poor focus, producing a low parasite staining; so, the U-Net model cannot segment it correctly.
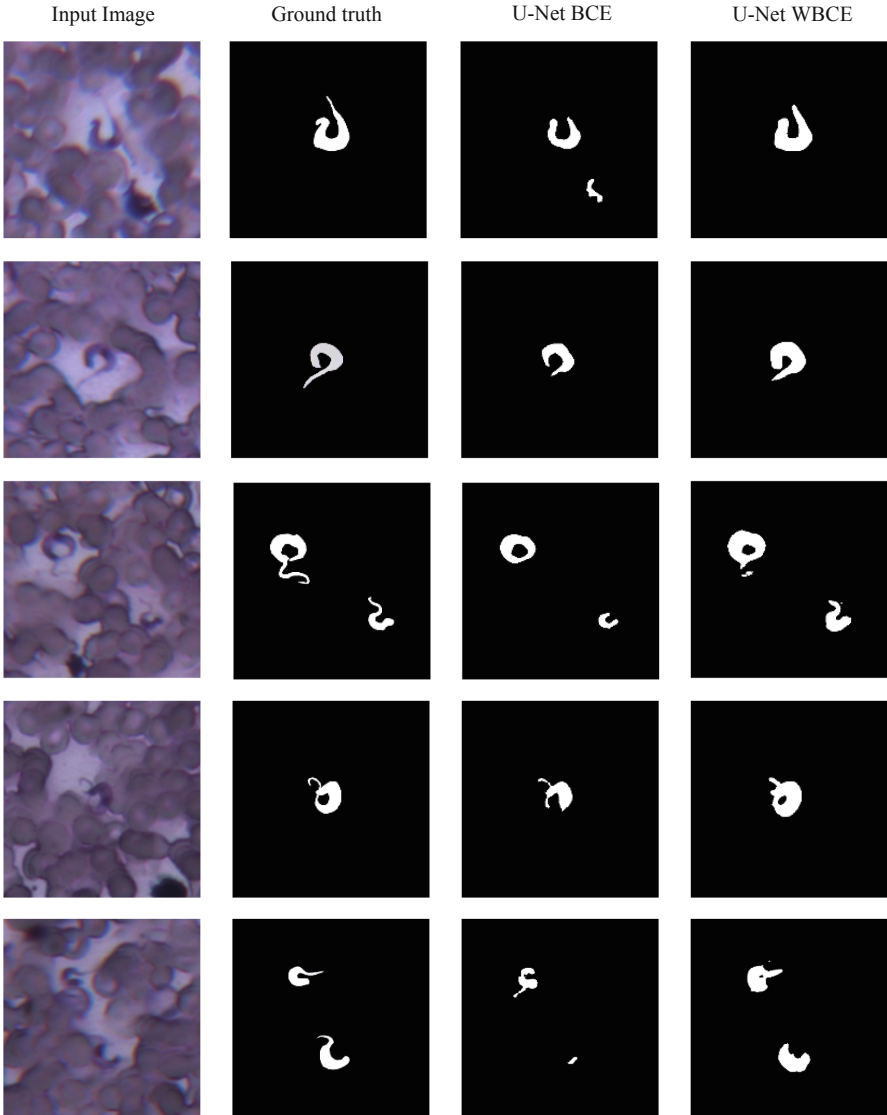
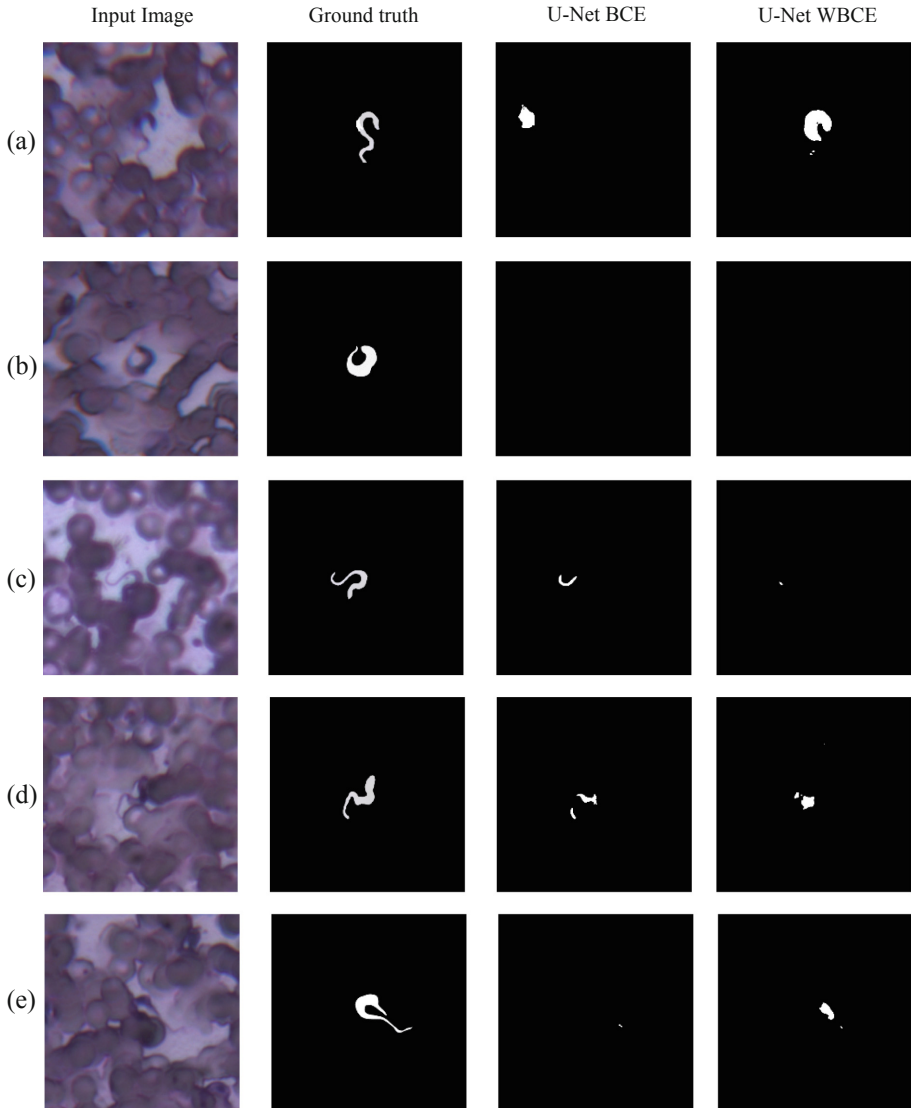**Fig. 8.** Qualitative results of predictions returned by the U-Net

Fig. 9. Qualitative results of wrong predictions returned by the U-Net

6   Conclusions

We proposed an automated method based on deep learning to segment the *T. cruzi* parasite on blood samples images as an alternative to a tired manual visual inspection with a microscope. A proper segmentation can assist experts to easily confirm if the *T. cruzi* parasite is present on the images to get a quick diagnosis of the Chagas disease, saving time and getting relevant information for further use. We chose the U-Net model because it is commonly used for biomedical image segmentation showing satisfactory

results. We conduct two experiments with the U-Net model. The first experiment consists on training the U-Net model with the Weighted Binary Cross-Entropy Loss function to reduce the class imbalance between *T. cruzi* parasite class and background class. The second experiment consists on training with the regular Binary Cross-Entropy Loss. We get the best F2 score and recall value on the first experiment. We report an F2 value of 0.8013, a recall value of 0.8702, a precision value of 0.6304 and a Dice score value of 0.6825. In comparison with the state of the art methods, our proposal obtains good results without executing a pre-processing of the input data (as the definition and feature extraction) or a post-processing of the results, obtaining a quick and reliable model to segment the *T. cruzi* parasite. Qualitative results show a good segmentation performance for *T. cruzi* parasite. With these promising results, we have shown that a deep learning model based on the U-Net architecture can achieve a proper *T. cruzi* parasite segmentation in blood sample images.

References

1. W. H. Organization Homepage: Chagas disease American Trypanosomiasis. https://www.who.int/en/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis). Accessed 02 Apr 2019
2. W. H. Organization: Chagas disease American trypanosomiasis. https://www.who.int/chagas/disease/en/. Accessed 22 May 2019
3. Uc-Cetina, V., Brito-Loeza, C., Ruiz-Piña, H.: Chagas parasite detection in blood images using AdaBoost. Comput. Math. Methods Med. **2015** (2015). 13 pages. https://doi.org/10.1155/2015/139681
4. Uc-Cetina, V., Brito-Loeza, C., Ruiz-Piña, H.: Chagas parasites detection through Gaussian discriminant analysis. Abstraction Appl. **8**, 6–17 (2013)
5. Delas Peñas, K., Rivera, P., Naval, P.: Malaria parasite detection and species identification on thin blood smears using a convolutional neural network, pp. 1–6 (2017). https://doi.org/10.1109/chase.2017.51
6. Poostchi, M., Silamut, K., Maude, R., Jaeger, S., Thoma, G.: Image analysis and machine learning for detecting malaria. Transl. Res. **194**, 36–55 (2018)
7. Mehanian, C., et al.: Computer-automated malaria diagnosis and quantitation using convolutional neural networks. In: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), pp. 116–125 (2017). https://doi.org/10.1109/ICCVW.2017.22
8. Soberanis-Mukul, R., Uc-Cetina, V., Brito-Loeza, C., Ruiz-Piña, H.: An automatic algorithm for the detection of Trypanosoma cruzi parasites in blood sample images. Comput. Methods Prog. Biomed. **112**, 633–639 (2013)
9. Górriz, M., Aparicio, A., Raventós, B., Vilaplana, V., Sayrol, E., López-Codina, D.: Leishmaniasis parasite segmentation and classification using deep learning. In: Perales, F.J., Kittler, J. (eds.) AMDO 2018. LNCS, vol. 10945, pp. 53–62. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94544-6_6
10. Latif, J., Xiao, C., Imran, A., Tu, S.: Medical imaging using machine learning and deep learning algorithms: a review. In: 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), pp. 1–5 (2019). https://doi.org/10.1109/icomet.2019.8673502

11. Ponce, C., et al.: Validation of a rapid and reliable test for diagnosis of Chagas' disease by detection of Trypanosoma cruzi-specific antibodies in blood of donors and patients in Central America. J. Clin. Microbiol. **43**, 5065–5068 (2005)

12. Egüez, K., et al.: Rapid diagnostic tests duo as alternative to conventional serological assays for conclusive Chagas disease diagnosis. PLoS Neglect. Trop. Dis. **11**(4), (2017). https://doi.org/10.1371/journal.pntd.0005501

13. Soberanis-Mukul, R.: Algoritmos de segmentación de trypanosoma cruzi en imagenes de muestras sanguineas. Master's thesis, Universidad Autónoma de Yucatán (2014)

14. Soberanis-Mukul, R.: Detección de trypanosoma cruzi en imágenes obtenidas a partir de muestras sanguíneas. Bachelor's thesis, Universidad Autónoma de Yucatán (2012)

15. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, William M., Frangi, Alejandro F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

16. Losses for segmentation. https://lars76.github.io/neural-networks/object-detection/losses-for-segmentation/. Accessed 03 Sep 2019

17. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: IEEE International Conference on Computer Vision (ICCV 2015), vol. 1502 (2015). https://doi.org/10.1109/iccv.2015.123

18. Google Colaboratory. https://colab.research.google.com. Accessed 15 Oct 2019

19. What is the F2 score in machine learning? https://www.quora.com/What-is-the-F2-score-in-machine-learning. Accessed 16 Sep 2019