**Introduction:**

Chagas disease, also known as American trypanosomiasis, is a chronic infection caused by Trypanosoma cruzi, a protozoan parasite that affects millions of individuals worldwide and it’s native to Latin America, according to epidemiological studies. The size of the parasite varies as thin shapes are nearly 20 µm in length and 1 µm in width whereas the thick shapes are about 15 µm and 4 µm in length and width respectively [2]. More than 25 million individuals are at danger of contracting the disease, and around 10 million people have already been infected, primarily in Latin America. Chagas disease is responsible for around 20,000 fatalities per year [2]. Diagnosis in the acute phase is critical since it promises for cure if treatment is initiated [3]. Currently, the detection of the chagas parasite is conventionally performed via microscope examination of the peripheral blood smears and is performed by skilled medical professionals, but the cumbersome and time-consuming nature of the process, requires automizing the method. Furthermore, high number of cases, limited skilled professionals and non-standard measures of accuracy due to varying operators’ judgements are other limitations of the manual process. Thus, machine learning and deep learning based automized methods are required for fast, minimized efforts, and reliable detection of the disease.

**Related Work:**

Even though Chagas disease is a potentially fatal illness for the worldwide population, only a few machine learning models for automatically detecting or segmenting Trypanosoma cruzi in blood sample have been proposed [11−14]. In 2013, Cetina et al. [11] have proposed a Gaussian discriminant analysis method to detect T. cruzi parasite that achieved a mean accuracy of 91 percent; however, its performance was heavily reliant on parasite image color features, implying that any other blood structure with similar features could affect the algorithm's performance. Soberanis-Mukul [13] has used multiple algorithms to create a technique for segmenting T. cruzi parasites based on picture super-pixels. Even though its Gaussian and SVM classifiers produced the highest performance results, this technique produces a high number of false positives and does not retain parasite shape. [9]A binary classifier with K-nearest neighbors is trained, with sensitivity and specificity of 0.98 and 0.80, respectively.In [3], the authors applied Adaboost with support vector machines as classifier, and reported 100% sensitivity and 93.3% specificity (i.e. precision) .But their methods are time consuming and require manual work due to steps like preprocessing images, extracting features, selecting features and designing haar-cascade specific for chagas parasite. Deep learning based models don’t need such efforts as it automatically learns and selects the relevant features. Recent years, Deep learning and CNN based models have showed great promise to solve medical image analysis and revolutionized computer vision tasks, with an emphasis on object recognition [7]. They[ ] proposed Pre-trained MobileNetV2 based deep learning model for feature extraction followed by fine-tuned dense layer of single-neuron for automated T. cruzi detection in microscopy blood tests. They achieved 96.4% accuracy on validation dataset and the precision, sensitivity, F1-score values were 95.4%, 97.6% and 96.5% respectively. But their test accuracy was 72% which they suggested was an over trained issue and low data size issue[]. This reflects a doubt in the effectiveness of their model for real world scenario. We proposed and compared different models which gave much higher classification accuracy for training, validation, and test dataset. In the medical imaging field, CNNs have been extensively used to assist in the diagnose of several diseases and to improve clinical decisions [15, 16]. This encouraged us to apply CNN and deep learning based architectures to effectively classify parasite which could aid Chagas disease diagnosis, particularly in non-endemic locations where illness has emerged as a result of migratory movement [4] and might also aid blood donor triage when demand spikes, as in the case of blood banks [2].

**Methodology:**

We have trained, fine-tuned, and compared various deep learning and machine learning methods to classify chagas parasite. Some of models were pre-trained on ImageNet dataset and then fine tunned on our chagas data set whereas other models were only trained on chagas images.

**Res2\_SepConvNet\_SVM Model:**

We have proposed Res2Net followed by Separable Convolutional block to extract out features from the images and then used 3 Fully connected layers of 512 , 128 and 64 nodes to reduce the extracted features. At last, we added SVM based Classification layer to classify our images and achieved Validation Accuracy as 98.67% with precision and recall as 98.21 and 1.00 respectively. Whereas this model showed 98% accuracy on Test Data having precision and recall as 97.61% and 98.95% respectively.

There are two types of residual blocks used in the Res2Unet architecture , one is called “**Res\_Block**” and it takes the input and then applies two Convolution block on it which is followed by Lambda function. Then it applies one Convolution block, having kernel size of 1\* 1 , on the input . After that it takes these two outputs from Convolution blocks and then adds them which acts as the output of the “**Res\_Block”**.

**Res\_Block**

**Output**

**Input**

**Add**

**Lambda Function(0.5\*It’s input)**

**Conv Block2**

**Conv Block1**

**Conv Block (1\*1 Kernel)**

**Res\_Block2:**

**Output**

**Input 1**

**Res\_Block**

**Average**

**Input 2**

Another residual block is called “**res\_block2**” and it takes two inputs, one is a previous layer output, and another is an input of the previous layer . It takes the output of the previous layers and then applies two Convolution block on it. Then it applies one Convolution block, having kernel size of 1\* 1 , on the previous layer output . Then adds the output of these two Convolution Blocks. At last, it applies average to the output of this add layer and the 2nd input of this “**res\_block2”.** This output of the average layer serves as the output of the “**res\_block2”**.

**SepConvNet Block:**

A depth-wise spatial convolution, which works on each input channel independently, is performed first followed by a pointwise convolution, which adds the resulting output channels. The benefit of Separable Conv2D is that it requires less computation and it also computes the depth wise information in images which helps in classifying images.

**Output**

**MaxPoolLayer**

**Batch-Norm Layer**

**Separable Conv2d**

**Separable Conv2d**

**Input**

**SVM Layer:**

Instead of using a Dense layer with one node as classifier. We have used the Dense Layer which acts as an SVM classifier, for that the loss function was set as “hinge” and L2 was selected as a kernel regularizer. This SVM layer gave much better accuracy than using fully connected layer as classifier.

Among all classification models , this model performed best and involved two “**res\_block**” , two “**res\_block2**” followed by Separable Convolution block. Then we added 3 fully connected layers to reduce the feature vector, at the end, the last dense layer was added which acted as an SVM Classifier to classify the parasite images.

**Res\_Block1**

**INPUT**

**OUTPUT**

**SVM based\_ Dense\_ Layer**

**3 Dense Layers**

**Sep\_Conv\_Block**

**Res\_Block2**

**Res\_Block1**

**Res\_Block2**

**VGGNet Based Fine-tuned Model:**

We used VGGNet for two purposes, at first as a classifier and then as a Feature extractor followed by machine learning models for classification.

For the first technique, we fine-tuned the VGGNet pre-trained model on our dataset , after removing the fully connected layers from its architecture. Further, we added GlobalAveragePooling2D layer which reduced the features from 32768 to 512. At last, we added 4 fully connected layers having nodes as 512 , 128 , 64 and 1 respectively to classify the images. It gave 97.33% accuracy for Validation Data and 97.33% accuracy for Test data.

For second technique, VGGNet was used as a feature extractor , PCA for feature reduction and Machine learning models, like SVM, XGBoost and Random Forest, acted as Classifier which gave a little better accuracy then the above described VGGNet based classifier and had less computational cost, thus we introduced this Model in our work as well.

Initially, we took the Pretrained VGGNet Model, fine-tuned on our dataset, and then extracted 32768 features from it. Then we applied Standard scaling to normalize the data and PCA technique to reduce features to only 125 features, which explained 90% of the variance of the features. Further, we used these reduced features for classification using XGBoost, Random Forest and SVM models and observed 98% test accuracy for each of them. In the next step, we gave all 32768 extracted features , from fine-tunned VGGNet, to the models but their performance was decayed to 88.6 , 96 and 97.33 percentage test accuracy for SVM , XGBoost and Random Forest respectively.

**XCEPTION Based Model:**

XCEPTION Model consist of Parts called Entry Flow, Middle Flow and Exit Flow.

Entry Flow has two Convolution Block of 32 and 64 channels respectively, in the starting, and then followed by 3 blocks, each consist of 2 Separable Convolution block with One MaxPooling Layer. The output of 1st two Convolution block is added to the output of each of these three blocks.

Middle Flow consist of 3 Separable Convolution Blocks and a residual connection which carries the input and adds it to the output of these 3 Separable Convolution Blocks. The Middle Flow can be repeated any number of times but in the original paper[ ] it was repeated 8 times. But I modified the architecture as this huge architecture fits better with highly complex scenarios and tasks. We have selected only 2 repetitions of Middle Flow block as our task performed better with this architecture and increased from 94% to 96% accuracy for Test data. Other changes that we performed were: Before dense layers, we have removed Separable Convolution block and these changes have made the architecture less complex and reduced the parameters from 88 million to 58 million and thus, computational it’s less costly, performs better and doesn’t overfit our dataset.

Exit block has 3 blocks, each consist of 2 Separable Convolution block with One MaxPooling Layer. The input to exit flow, after being transformed by Convolution Block, is added to the output of these three blocks.

Diagram

Description automatically generatedTable

Description automatically generated with medium confidence

We have tried other models based like Separable Conv Based Model, Res2Net based Model, ResNet50 Model with transfer learning , Inception Based Fine-tuned Model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MODEL** | **Test Accuracy** | **Validation Accuracy** | **Precision** | **Recall** |
| ResNet50 Transfer Learning | 95.33 | 94.00 | 93.17 | 99.00 |
| INCEPTION Fine-tuned Model | 98.00 | 96.67 | 96.28 | 100 |
| XCEPTION Based Modified Model | 96.00 | 98.67 | 95.36 | 97.84 |
| SeparableConvNet | 95.33 | 98.00 | 92.22 | 100 |
| Res2Net Based Model | 96.67 | 98.00 | 95.39 | 98.95 |
| VGGNet Model | 97.33 | 98.00 | 96.28 | 98.95 |
| VGGNet\_PCA\_SVM | 98.00 | 98.00 | 97.00 | 100 |
| VGGNet\_PCA\_XGBOOST | 98.00 | 96.00 | 98.00 | 99.00 |
| VGGNet\_PCA\_Random Forest | 98.00 | 97.33 | 98.00 | 99.00 |
| Res2\_SepConvNet\_SVM | **98.00** | **98.67** | 97.61 | 98.95 |

**Metrics:**

**Precision**: Out of all the positive class predications made by the model, how many were actually belonging to positive class is measured by Precision. In other words, Precision is the percentage of true positive examples among the positive examples classified by the model.

Text

Description automatically generated with low confidence

**Recall** : The fraction of positive examples labeled as positive among the total number of positive examples is known as recall, also known as sensitivity.

A picture containing text

Description automatically generated

**F1 Score:** The F-score, also known as the F1-score, is a metric for how accurate a model is on a given dataset**.** It is helpful when we deal with unbalanced samples.



**Accuracy:** The number of correctly classified examples divided by the total number of examples is the definition of accuracy. Although accuracy is useful, it does not account for the intricacies of class imbalances.



**Experimental Work and Results:**

Initially, we have introduced Res2Net based Model consisting of 5 blocks with Res\_block1 and Res\_block2 alternatively, to classify images and found the accuracy as 96.67 for test data which was comparatively less and then we implemented other models like XCEPTION based and SeparableConvNet Models which gave roughly same test accuracy as previous model . After that, we implemented VGGNet classifier followed by models like VGGNet\_PCA\_SVM , VGGNet\_PCA\_RF , VGGNet\_PCA\_XGBoost and found that very high test accuracy of 98% was obtained on test and validation data by SVM based VGGNet\_PCA\_SVM classifier so, we introduced a novel model based on the benefits of previous models and blocks and came up with Res2\_SepConvNet\_SVM which had benefits of Res2Net and SepConvNet blocks as a feature extractors which were added with DNN block of reduced nodes as 512, 128 and 64 for feature reduction and then SVM based dense layer for classification. At last, we were able to get highest accuracy on test data, 98.67% with high precision and recall as 97.61 and 98.95 percentage respectively.

We extracted features from VGGNet , as CNN models are good at learning feature mapping, further, to lessen the computation time for the classification Model, we reduced the features using PCA technique. At first, we extracted 50 features which represent 90% of the variance in the total features(32768) and feed them to the machine learning models. We found that SVM classifier performed better but other models , like Random Forest and XGBoost, didn’t gave better results. So, we reduced the features amount to 20 , 9 and 3 step by step and observed that XGBoost performed best at 20 features whereas XGBoost and SVM showed highest accuracy even at 9 and 3 features respectively. Other important point to consider is that Random Forest performed much better at 9 features compared to 32768 features. Similarly, XGBoost showed higher test accuracy, 98%, for lower feature representation of 20 compared to 32768 feature representations, having 97.33% as test accuracy. Other interesting finding was that same extracted features, from fine-tuned VGGNet, were feed to DNN block which was connected to VGGNet and were also feed to Machine learning models. The DNN consisted of one global averaging pooling followed by 3 fully connected layers of 512, 128, 64 and 1 node respectively. This architecture of reduced nodes was selected so that the model can train and find best feature representation possible before the classification is done at the last layer. But we found that Machine learning models performed better than this VGGNet with DNN block , as SVM and Random Forest showed 98% test accuracy at 9 feature representations while VGGNet with DNN block showed 97.33% test accuracy. To conclude, when right set of features were extracted and feed to models can extract much better accuracies. The possible reasons for such small accuracy difference might be other than this, but this can also be one possible reason. Other possible reasons can be less training of model or the higher complexity of the model.

**Conclusion:**

We have compared multiple models, some were fine-tuned on our dataset but most of these were trained from scratch on our dataset, and we obtained good accuracy on training , validation, and test images with high precision and recall except for ResNet50 and SeparableConvNet Models. According to our knowledge, this paper is the first one to compare various classification based deep learning and machine learning models on chagas parasite. This work can help in chagas disease computer-aided diagnosis, a fast and effective method, of detecting it in acute phase which can potentially save millions of lives. WHO considered chagas as an tropical neglected disease[] and emphasized a need for proper research and implementation to cure it which inspired us to contribute a little in this bigger objective. The accuracy of our best model was 98.67% and 98% on validation and test data respectively; however, we would like to increase the performance further in our future work. Possible work includes using more dataset by collecting it or introducing novel techniques of generating new and relevant images using GAN under supervision and approval of medical experts, combining predications or architecture of multiple models to reduce misclassification error, extracting most relevant features which can further reduce computational cost and increase accuracy, visualization of features in different layers to gain better understanding of features generation to improve the model architecture.