

A Novel Approach of Malaria Parasite Disease Detection Using Hybrid Deep learning Model

B S Panda¹, Mounika Pattnaik²

^{1,2}(Dept. of CSE, Raghu Engineering College, Visakhapatnam, AP, India)

Abstract: Anopheles mosquito bites from female Anopheles mosquitoes carry the Plasmodium parasites that cause malaria, a blood disease. The worst disease on the planet, malaria creates a lot of busy work for the health department. The conventional method of diagnosing malaria involves a graphic examination of human blood smears under a microscope by lab or trained workers for parasite-infected red blood cells. This method is ineffective, and the diagnosis is dependent on the knowledge and experience of the examiner. Prior to this, deep learning algorithms had been used to diagnose malaria blood smears for diagnosis before methods that use machine learning (ML) techniques and hand-engineered features on microscopic pictures of smears call for competence in analyzing textural, morphological and positional alterations of the region of interest (ROI). Mostly convolutional neural networks (CNNs), a subset of deep learning (DL) models, offer greater performance with end-to-end feature extraction and categorization. Therefore, automated malaria screening employing DL methods might be useful as a diagnostic tool. In order to improve disease screening, we examine the presentation of customized VGGNET+CNN based DL models as feature extractors for identifying the parasitized and uninfected cells. We use experiments to choose the best model layers for extracting features from the primary data. The utilization of customized CNNs as a potential technique for feature extraction for this purpose is shown by statistical validation of the actual results.

Keywords: Deep Learning, Convolutional Neural Networks, Machine Learning, Malaria, Feature extraction.

I. Introduction

The parasite Plasmodium causes malaria, a blood disease spread by mosquito bites, which is contracted by female Anopheles mosquitoes. Humans can become infected with a variety of parasites, including *P. falciparum*, *P. malariae*, *P. vivax*, and *P. ovale*, although *P. falciparum*'s consequences can be fatal. The World Health Organization (WHO) estimated that In 2016, there were 212 million cases of the condition worldwide. Microscopic thick and thin blood smear exams are the most popular and efficient method for illness diagnosis. Thick blood smears are used to either identify the kind of parasite causing an infection or to detect the presence of parasites (Centers for Disease Control and Prevention, 2012).

Large-scale diagnoses in areas where disease is common and resources are scarce may impose inter-observer variability and liability that may have a negative effect on the diagnostic accuracy, which heavily depends on human competence (Mitiku, Mengistu & Gelaw, 2003) [1]. Polymerase chain reaction (PCR) and rapid diagnostic tests (RDTs) are other approaches that are employed, although PCR analysis has performance limitations, and RDTs are less cost-effective in areas where disease is prevalent (Hawkes, Katsuva & Masumbuko, 2009) Meaningful feature representation is essential to the effectiveness of machine learning (ML) techniques used for medical data analysis in achieving desired outcomes.

The popular of computer-aided diagnosis (CADx) software that uses image analysis relies on ML approaches with manually created characteristics for making decisions (Ross et al., 2006; Das et al., 2013; Poostchi et al., 2018). Assessing these variances needs skill because the region of interest (ROI) on the images varies in size, backdrop, angle, and position. Deep Learning (DL), also known as deep hierarchical learning, is utilized with notable success to overcome difficulties in developing manually built features that detain differences in the original data (LeCun, Bengio & Hinton, 2015) [2].

In order to automatically find a cascade of layers of non-linear processing units are used by DL models to describe hierarchical feature representations in the raw data. Higher-level features are abstracted from lower-level features in order to aid in the understanding of challenging, non-linear decision-making processes, resulting in end-to-end feature extraction and categorization (Schmidhuber, 2015).

DL models are extremely scalable because, unlike kernel-based methods like Support Vector Machines (SVMs), they perform better as data size and processing resources increase (Srivastava et al., 2014). The spatial local correlation between the nearby pixels or voxels is a key source of information for images. By using the techniques of local receptive fields, shared weights, and pooling, Convolutional Neural Networks (CNN), a

family of DL models, are created to take advantage of this information (Hinton, E., Srivastava & Krizhevsky 2012) [3].

In 2012 the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) was won by Alex Krizhevsky's Alex Net, a CNN-based DL model that significantly improved CNN performance while classify natural photos (Krizhevsky, Sutskever & Hinton, 2012). In order to better the Inception-V3 model on the ImageNet (Deng et al., 2009) data classification task, several representative CNNs, including VGGNet (Simonyan & Zisserman, 2015), GoogLeNet (Szegedy et al., 2014), and ResNet (He et al., 2016), demonstrated depth-wise separable convolutions (Chollet, 2016). Densely Connected Convolutional Networks (DenseNet) is a CNN version that makes use of a network topology where each layer is directly connected to every subsequent layer (Huang et al., 2016) [4].

While employing substantially fewer parameters and calculations than the state-of-the-art, the model has produced notable advances. The availability of an enormous amount of annotated data goes hand in hand with the promising performance of CNNs. In order to help with visual recognition tasks due to the lack of Transfer Learning (TL) techniques are used to annotate medical images, where pre-trained DL models are either adjusted on the underlying data or used as feature extractors (Razavian et al., 2014) [5]. The general traits that these models have discovered from large datasets like ImageNet are applied to the task at hand.

It is more universal than situation-specific to apply previously learned skills to new situations. Since the publication of Razavian et al (2014) findings, it has been recognized that CNNs trained on enormous datasets can serve as feature extractors for a range of computer vision applications, resulting in improved performance over existing methods (Bousetouane & Morris, 2015).

In order to analyze movies that had a focus stack of the field of views of Leishman-stained slide images for the purpose of automating parasite detection, Gopakumar et al. (2018) [6] used a customized CNN model. Data collecting was carried out by the authors using a customized portable slide scanner and commercially available components, and they showed sensitivity and specificity of 97.06% and 98.50%, respectively. In conclusion, existing DL research have been assessed using randomized train/test splits and/or relatively limited image sets. No research has reported on the predictive models' efficacy at the patient level. Despite the positive reported results, cross-validation studies at the patient level are required to demonstrate the resilience of current techniques on a broader range of photos assessment at the patient level.

II. Literature Survey

As there is no information about bloom variations or other arte facts leaking into the training data, the photos in the independent test set represent completely hidden images for the preparation process, providing a more realistic performance evaluation of the predictive models. By doing this, prejudice and generalization errors would be decreased. Testing for statistically important performance differences would help in the process of choosing the best model before deployment. It is appropriate to point out that there is still much potential for improvement in this area even with the state-of-the-art.

In this previous study, we assessed how well pre-trained CNN-based DL models performed as feature extractors when it came to categorizing parasitized and misused cells to help with better infection screening. The following are some of this work's significant contributions: (a) presenting a evaluation of the effectiveness of pre-qualified and customized DL models as feature extractors for classifying parasitized and unused cells, (b) cross-validating the predictive model's performance at the patient level to reduce bias and generalization errors, (c) analyzing and choosing the best player in the pre-trained models to extract features from the underlying data, and (d) testing for the presence/absence of a statistician. The following paper is divided into three sections: "Materials and Procedures," which describes the materials and methods in more detail; "Results," which displays the results; and "Discussions and Conclusion," which explores the findings and wraps up the paper.

By utilizing ensemble techniques based on deep learning to automate the detection of the parasite using entire slide images of thin blood smears, our existing survey intends to create an effective system. This study will make use of the Snapshot Ensemble, a quick technique. Using the Snapshot Ensemble technique, several snapshots or weak learners can be produced at the expense of training just one neural network, which can then be integrated to create a single strong model. All experimental models were evaluated using the F1 score, Accuracy, Precision, Recall, Matthews Correlation Coefficient (MCC), Area under the Receiver Operating Characteristics (AUC-ROC), and Area under the Precision Recall Curve. (AUC- PR). Achieve an f1 score of 99.37%, precision of 99.52%, and recall of 99.23%, the snapshot ensemble model developed by integrating the snapshots of the EfficientNet-B0 pre-trained model surpassed all other models. The findings highlight the promise of model ensembles, which pool the predictive capacity of a number of weak learners to produce a single, powerful more effective design able to handle data from the actual world. Additionally, snapshot ensemble demonstrated effectiveness with a constrained training budget.

The Grad CAM experiment displayed the gradient activation maps of the last convolution layer to visually show where and what a model sees in an image in order to categorize it into a particular class. The models in this study accurately activate the stained parasite region of interest in the thin blood smear images. These pictures improve the model's dependability, explain ability, and transparency when AI-based models are implemented in the healthcare network.

The parasite Plasmodium causes malaria, a blood disease spread by mosquito bites, which is contracted by female Anopheles mosquitoes [7]. Humans can become infected with a variety of parasites, including *P. ovale*, *P. malariae*, *P. vivax*, and *P. falciparum*, although *P. falciparum*'s consequences can be fatal. In 2016, there were thought to be 212 million instances of the condition worldwide, according to the World Health Organization (WHO). Microscopic thick and thin blood smear exams are the most popular and efficient method for illness diagnosis. Thick blood smears are used to either identify the kind of parasite causing an infection or to detect the presence of parasites (Centers for Disease Control and Prevention, 2012). Large-scale diagnoses in areas where disease is common and resources are scarce may impose inter-observer variability and liability that may have a negative effect on the diagnostic accuracy, which heavily depends on human competence (Mitiku, Mengistu & Gelaw, 2003). Rapid diagnostic tests (RDTs) and polymerase chain reaction (PCR) are other approaches that are employed, although PCR analysis has performance limitations, and RDTs are less cost-effective in areas where disease is prevalent (Hawkes, Katsuva & Masumbuko, 2009) [8] Meaningful feature representation is essential to the effectiveness of machine learning (ML) techniques [9] used for medical data analysis in achieving desired outcomes.

III. Proposed Methodology

In order to extract the features for malaria disease identification, we suggested combining transfer learning (VGG16) with CNN. The judgment as to whether a specific cell is infected or healthy is where the difficulty begins start the machine's training by providing all of the image's properties. Therefore, a total of 27,558 images were gathered by gathering as many through the internet.

A. Data Collection:

To reduce the workload on traditional light microscopes in resource-limited areas and improve diagnostic precision, researchers at the Lister Hill National Center for Biomedical Communications (LHNCBC), a division of the National Library of Medicine (NLM), have developed a mobile application that runs on a standard Android device (Poostchi et al., 2018) [10] Giemsa-stained thin blood smear slides from 150 *p. falciparum*-infected and 50 healthy patients were collected and photographed at Chittagong Medical College Hospital in Bangladesh. For each little field of vision, slides were captured using the smartphone's built-in camera. The photographs were painstakingly described by a qualified slide reader at the Mahidol-Oxford Tropical Medicine Research Unit in Bangkok, Thailand. The de-identified images and annotations are stored in a database at NLM (IRB#12972), to recognize and separate the red blood cells, we used a level-set based algorithm (Ersoy et al., 2012) [11].

B. Cross-validation studies

There are equal numbers of parasitized and uninfected cells in the dataset's 27,558 photos of individual cells. Plasmodium was present in positive samples, while Plasmodium was absent in negative samples, although other items such as staining artifacts or contaminants were present. Five-fold cross-validation was used to evaluate the predictive models. To eliminate model biasing and generalization mistakes, cross-validation has been done at the patient level. The figure displays the number of cells for the various folds.

We have two folder images: 1. Parasitized 2. Uninfected

2757 pictures were taken, both of infected and parasitized subjects.

To meet the input necessities of a tailored and pre-trained CNN, the photos were re-sampled to 100 x 100, 224 x 224, 227 x 227, and 299 x 299 pixel resolutions, and normalized to aid in closer meeting. The models were trained on Google Colab's Tensor Flow platform using keras and Python.

C. Customized model configuration:

We also assessed a tailored, sequential CNN's performance in the job of distinguishing parasitized and uninfected cells for disease screening. Here suggest a sequential CNN that is comparable to the design favored by LeCun & Bengio (1995) for picture classification [12].

The proposed CNN consists of three convolutional layers and two fully linked layers segmented cells with a resolution of 1001003 pixels are the model's input. Utilizing 3 3 filters with 2 pixel sweeps, the

convolutional layers. In addition to having 64 filters in the third convolutional layer, the first and second convolutional layers each include 32 filters. The learning process is improved by the sandwich construction of convolutional/rectified linear units (ReLU) and suitable weight initialization (Rajaraman, Sivaramakrishnan, et al 2018) [13].

Convolutional layers are followed by max-pooling layers, which summarize the outputs of neighbouring neural groups in the feature maps with a pooling window of 2 2 and 2 pixel strides. The first fully-connected layer, which comprises 64 neurons, receives the pooled output of the third convolutional layer, while the second fully-connected layer feeds the SoftMax classifier. On the outputs of the first fully connected layer, dropout regularization (Srivastava et al., 2014) with a dropout ratio of 0.5 is used. Using stochastic gradient descent (SGD) and Nesterov's momentum, the multinomial logistic regression objective is optimized to train the model (LeCun, Bengio, and Hinton, 2015). (Botev, Lever and Barber. 2017). A randomized grid search technique is used to optimize the customized model for hyper-parameters (Bergstra & Bengio, 2012) [14]. We define the initial search ranges of the learning rate, SGD, and L2-regularization parameters as [1e7 5e2], [0.8 0.99], and [1e10 1e2], respectively. We evaluated the process. (Lipton, Elkan, and Naryanaswamy, 2014) [15] The customized model's Area under Curve (AUC), sensitivity, specificity, F1-score, and Matthews correlation coefficient (MCC).

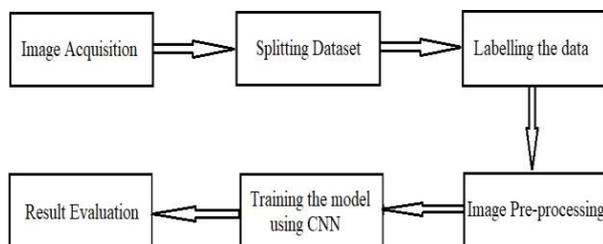


Fig.1: Architecture of CNN Model

D. Pre-trained and Feature extraction models:

We assessed the effectiveness of pre-trained CNNs including Alex Net (ILSVRC 2012 winner), VGG-16 (ILSVRC 2014 winner of the localization task), Xception, ResNet-50 (ILSVRC 2015 winner), and DenseNet-121 (ILSVRC 2015 winner of the best paper award) at extracting features from parasitized and uninfected cells. By using the randomized grid search method, the techniques were optimized for hyper-parameters. We set the initial search ranges for the learning rate, Nesterov's SGD, and L2-regularization parameters to be [1e5 5e2], [0.8 0.99], and [1e10 1e2], respectively. On top of the retrieved features, we trained a fully-connected form with dropout (the ratio of dropout 0.5) by instantiating the convolutional portion of the pre-trained CNNs. To aid in better classification, we empirically identified the best layer for feature extraction [16]. We assessed how well the trained CNNs performed in terms of correctness, AUC, sensitivity, specificity, F1-score, and MCC. The pre-trained CNNs' model architecture and weights were acquired from GitHub repositories.

E. Results

Cell detection and segmentation: As seen in fig.2, we used a level-set based model to identify and part the RBCs. We used a multi-scale Laplacian of Gaussian (LOG) filter to notice the centroids of each RBCs in the first step, which is cell detection. The cells are divided using the generated in order to keep the evolving contour inside the cell border; markers are placed within a level-set active contour framework. Utilizing average cell size, morphology opening techniques are used as post-processing to eliminate falsely discovered objects like staining artifacts. Since WBCs are not, white blood cells are detached utilizing a conversation correlation imposed on cell ground truth annotations.

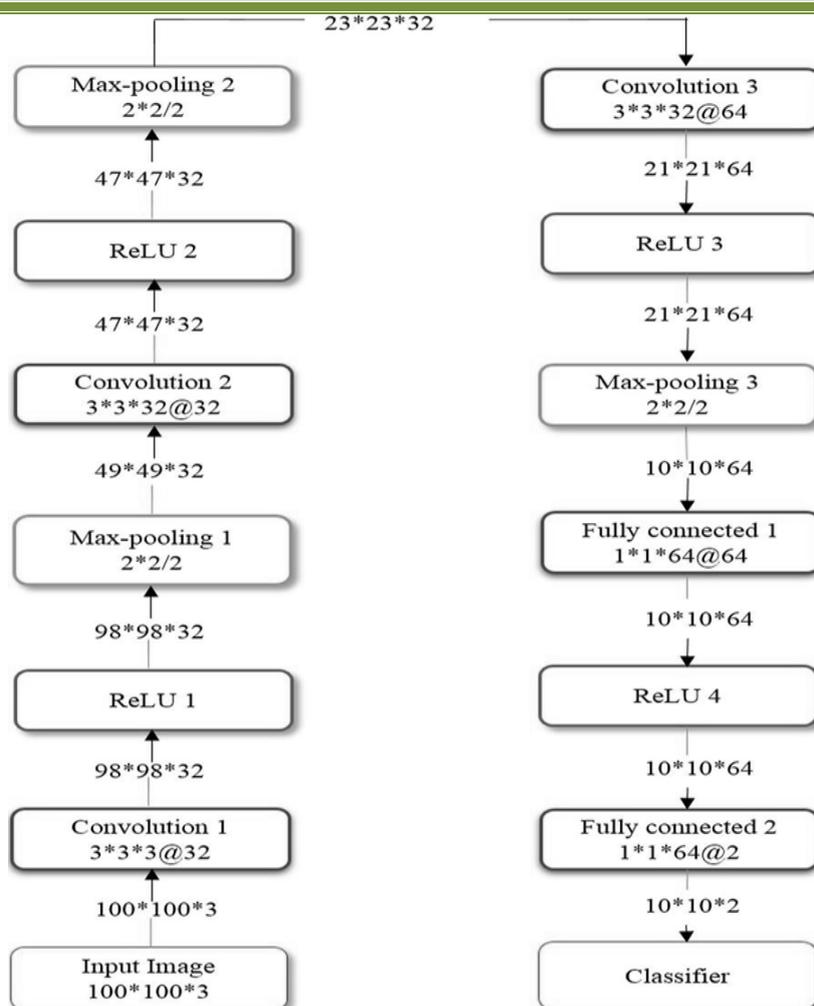


Fig.2: Feature Extraction of trained model

Actual/Predicted	Healthy-Cells	Malaria-Cells
Healthy-Cells	4040	78
Malaria-Cells	268	3993

Table1: The Fine-Tuned CNN (CNN+VGGNET) confusion matrix

CNN	Acc	Class.	Sen	Pre	F1	Fb	Spe	Fal	Mat
Fine -Tuned CNN	0.96	0.03	0.93	0.98	0.96	0.96	0.98	0.01	0.92

Table2: The results of the three convolutional neural networks calculation

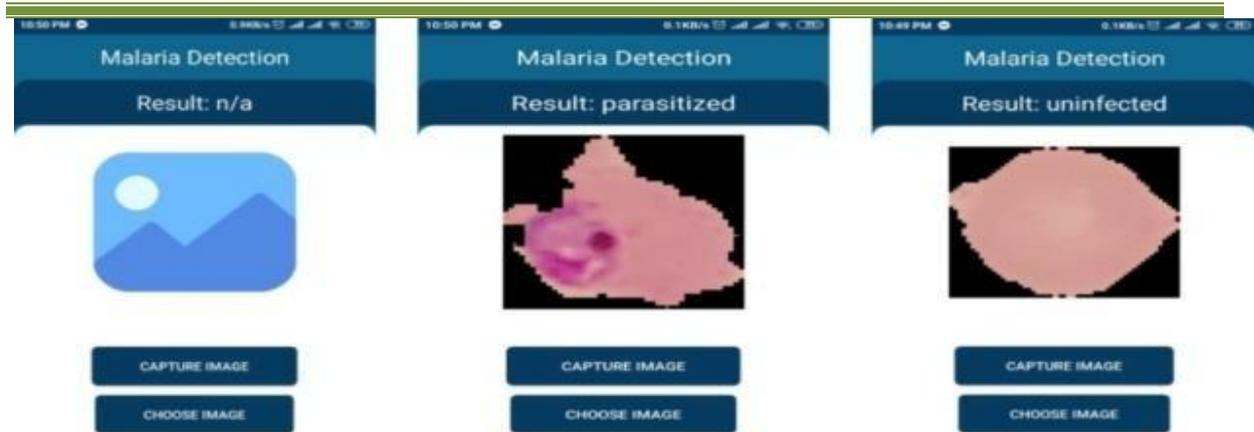


Fig. 3: Web Application for malaria detection

IV. Conclusion & Future Scope

Due to hyper-parameter tuning, implicit regularization imposed by reduced convolutional filter sizes, and aggressive dropouts in the fully connected layers, the tailored model was able to converge to an ideal outcome. L2 regularization was used to lessen the impact of model overfitting and provide a more effective result. Pre trained CNNs by extracting features from various layers in order to determine the best layer from the underlying data for feature extraction.

We noticed that the presentation of the layer prior to the arrangement layer was worse for the pre-trained CNNs than it was for the other layers. The combined results showed that when it came to distinguishing parasitized and uninfected cells, the VGGNET+CNN substantially outperformed the other better accuracy models under evaluation. We found that the pre-trained VGGNET+CNN achieved the maximum accuracy when doing Kruskal-Wallis H and post-hoc analyses.

References

- [1] Mitiku, K., Mengistu, G., & Gelaw, B. (2003), The reliability of blood film examination for malaria at the peripheral health unit. *Ethiopian Journal of Health Development*, 17(3), 197-204.
- [2] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521.7553 (2015): 436-444.
- [3] Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R. (2012), improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*.
- [4] Silver, David, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser et al. (2016), "Mastering the game of Go with deep neural networks and tree search." *nature* 529, no. 7587, 484-489.
- [5] Sharif Razavian, Ali, et al. (2014), "CNN features off-the-shelf: an astounding baseline for recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*.
- [6] Gopakumar, Gopalakrishna Pillai, et al. (2018), "Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner." *Journal of biophotonics* 11.3 e201700003.
- [7] B. S. Panda, S. S. Gantayat, Ashok Misra, (2014), "Rough Set Rule Based Technique for the Retrieval of Missing Data in Malaria Diseases Diagnosis" In: *SPRINGER FMB Series*, ISBN: 978-981-287-260-9.
- [8] Hawkes M, Katsuva J, Masumbuko C. (2009), Use and limitations of malaria rapid diagnostic testing by community health workers in the war-torn Democratic Republic of Congo. *Malaria Journal* 8(1): 308 DOI 10.1186/1475-2875-8-308.
- [9] Bergstra J, Bengio Y. (2012), Random search for hyper-parameter optimization. *Journal of Machine Learning Research* 13: 281–305.
- [10] Poostchi, Mahdih, et al. "Image analysis and machine learning for detecting malaria. (2018), " *Translational Research* 194, 36-55.
- [11] Ersoy I, Bunyak F, Higgins J, Palaniappan K. (2012), Coupled edge profile geodesic active contours for red blood cell flow analysis. In: *Proceedings of the 9th IEEE international symposium on biomedical imaging*, 2–5 May, Barcelona, Spain. Piscataway: IEEE, 748–751 DOI: 10.1109/ISBI.2012.6235656.
- [12] LeCun, Yann, and Yoshua Bengio. "Convolutional networks for images, speech, and time series." *The handbook of brain theory and neural networks* 3361.10 (1995): 1995.

- [13] Rajaraman, Sivaramakrishnan, et al. (2018), "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images." PeerJ 6 e4568.
- [14] Bergstra, James, and Yoshua Bengio. (2012), "Random search for hyper-parameter optimization." Journal of machine learning research 13.2.
- [15] Lipton, Zachary Chase, Charles Elkan, and Balakrishnan Narayanaswamy. (2014), "Thresholding classifiers to maximize F1 score." arXiv preprint arXiv: 1402.1892.
- [16] Dong Y, Jiang Z, Shen H, David Pan W, Williams LA, Reddy VVB, Benjamin WH, Bryan AW. (2017), Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells in: 2017 IEEE EMBS international conference on biomedical and health informatics, BHI 2017. Piscataway: IEEE, 101–104 DOI: 10.1109/BHI.2017.7897215.