Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 189–198

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Identifying Colorectal Tumor For Single Cell RNA Sequence Using Rectified Linear Unit With Stochastic Gradient Descent

Mothe Rajesh[a], Biswajit Senapati[a,b,*], Ranjita Das[b], Sheshikala Martha[a]

[a]*SR University, Warangal, Telangana and 506371, India*
[b]*National Institution of Technology Mizoram, Chaltlang, Aizawl, Mizoram and 796012, India*

Abstract

To detect cell clusters in whole human colorectal tumor cells, mechanisms based on single-cell RNA sequencing have been reported. To address such issues, a deep learning technique for single-cell analysis was recently developed, with promising results. Single-cell RNA sequence data contains detailed information about the transcriptome profile. Through gene expression, this information is presented in the form of patterns. The neural network is shown here to help understand these data representatives. Then we use the differential expression to identify distinct cell types and gene markers. Finally, the proposed identifying colorectal tumor for single-cell RNA sequencing that Rectified Linear Unit (ReLu) with Stochastic Gradient Descent(SGD) is compared to other recently developed models such as ReLu with Adam, ReLu with Limited-BFGS-B, TanH with Adam, TanH with SGD, TanH with Limited-BFGS-B, Sigmoid with Adam, Sigmoid with SGD, Sigmoid with Limited-BFGS-B, Linear with Adam, Linear with Limited-BFGS-B and Linear with SGD. At the moment, the results from identifying colorectal tumor for single-cell RNA sequencing that ReLu with SGD performs better than other recently developed models.

Keywords: Preprocessing, Hidden layer, Rectified linear , Stochastic gradient decent , Regularization;

1. Introduction

Single-cell RNA sequencing (scRNA-seq) has emerged as the most valuable and widely used technique in research. It has been thoroughly evaluated for a wide range of biological research. The transcriptome profiles for bulk cell sequencing and single-cell sequencing are compared. Every cell in bulk cell sequencing is combined to form a single entity, but it is extremely difficult to identify each cell individually. Because the data in the bulk sequence is mixed together and presented in a homogeneous form, it is difficult to find every feature separately from the bulk sequence. As a result, we developed scRNA-seq, which can solve problems associated with homogeneous systems by identifying individual cells[1]. Then deep learning algorithms rely on neural networks, which are a subset of data science. One of the decisions we will make when building a neural network is which activation function to use within the network's

* Corresponding author. Tel.: +91-828-046-5095 ; fax: +0-0-0-0.
E-mail address: biswajitkjbj@gmail.com

10.1016/j.procs.2023.01.001

hidden and output layers[2]. The activation function calculates the weighted total and includes extra bias to determine whether or not to stimulate a neuron. A neural network is just a linear regression model without an activation function. The activation function changes the data in a nonlinear style, and it will evaluate and perform more difficult tasks[3].

1.1. Activation functions

The most commonly used activation functions are briefly discussed in the below sections.

- **Linear:** This activation function doesn't change the weights of input in a neural network, instead, it throws the weights directly. Due to this behavior, it is also called a linear activation function [4]. One of the most significant limitations of a linear activation function is that it doesn't work with back propagation. If it is used, the entire neural network will collapse. Another issue with this type of function is that it assumes the first layer of the network is constant and that it will remain a linear function as the number of layers increases. It can be represented mathematically as
$$f(x) = x \tag{1}$$
- **Logistic:** This function is also called a sigmoid function. As an input, it takes any real value and gives the resulted values ranging from 0 to 1. If the input value is more positive then the output is near 1, whereas the lower input value means the output is near 0 [5]. This logistic activation function is commonly used. It can be represented mathematically as
$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$
- **Tangent hyperbolic(Tanh):** This function has the most similar behavior to the sigmoid activation function. As an input, it takes any real value and gives the resulted values ranging from -1 to 1. If the input value is more positive then the output is near 1, whereas the lower input value means the output is near -1. The Tangent hyperbolic function has some advantages, such as the fact that its output is zero-centered, allowing us to consider the output as highly negative, neutral, or highly positive [6]. It can be represented mathematically as
$$f(x) = \frac{(e^{x}-e^{-x})}{(e^{x}+e^{-x})} \tag{3}$$
- **Rectified linear unit(ReLu):** It is the most commonly used function among all the others. In this function all the neurons will not activate at once, this is the major advantage of it. Because of these characteristics of it, if the input is negative then it will convert it into zero as a result the neuron will not be activated [7]. The representation of the ReLu function is mathematically mentioned below:
$$f(r) = max(0, r) \tag{4}$$

1.2. Optimization algorithms

Deep learning algorithms frequently rely on the optimization of some objective function, selecting an optimization algorithm is an important part of the learning process. We will look at three of the most commonly used optimization algorithms in deep learning problems: SGD, L-BFGS, and Adam. This section contains a description of the algorithms mentioned.

- **Adaptive Moment Estimation(Adam):** Adam consists of two methodologies, one is momentum and the other one is a root mean square propagation. To design any model using this optimizer requires less computational time and less memory. It has first-order and second-order momentum to measure. The equations for this momentum is given below[8].
$$P_t = \beta_1 * P_{t-1} - (1-\beta_1) * r_t \tag{5}$$
$$Q_t = \beta_2 * Q_{t-1} - (1-\beta_2) * r_t^2 \tag{6}$$
- **L-BFGS:** Quasi-Newton methods are algorithms that seek to achieve some of the benefits of Newton's method without the computational complexity of the hessian. The main idea is to approximate the hessian in a good

way to avoid explicitly calculating all of the objective function's second derivatives. Because second-order methods have a higher computational burden, they are less common (in deep learning) than first-order methods. However, we will look at one algorithm from this family: the BFGS algorithm [9]. The BFGS algorithm requires storing the previous matrix for the next iteration, which can severely limit the problems that can be approached by the algorithm. When training neural networks, it is common to have millions of weights and biases, making it impossible to store the matrix between iterations. The L-BFGS algorithm approximates the BFGS algorithm while requiring less computer memory, making it a promising neural network algorithm.

- **Stochastic Gradient Descent(SGD)**: To optimize neural networks, it is the most widely used algorithm among the L-BFGS and adam [10].To address the drawbacks of L-BFGS and adam, SGD is one of the best choices in deep neural network. Based on how much data is utilized to find the gradient of the objective function, different types of variants are categorized.

$$R_{t+1} = R_t - \gamma_t \nabla_r S(u_t - r_t) \tag{7}$$

The primary contributions of this paper:

- A new model has been introduced by using a stochastic gradient descent algorithm, rectified linear unit activation function, hidden layer, and regularization which can analyze the single-cell RNA sequencing.
- For evaluating the performance measurement accuracy, precision, recall, and F1 types of parameters are included which are giving the best values for the designed model than the other recently developed models.
- At the end, identifying different types of cell types and gene markers present inside the colorectal tumor and normal mucosa.

The rest of this paper is organized as follows: section 2 is briefly given about a literature review and section 3 is about the proposed methodology in detail. In section 4, results and discussion detail information is availabled. Conclusion and future work is discussed in the section 5. In the last acknowledgement and references are presented.

2. Literature Review

On the hurricane harvey dataset, the authors[11] presented a stochastic gradient descent optimization technique using a corrected linear unit function. As a result, the performance accuracy is 95%, the recall is 96%, the precision is 97%, and the F1 is 96%. On the local medical center dataset, [12] proposed a stochastic gradient descent optimization technique with a sigmoid function. As a consequence, the performance accuracy is 81.13%, the recall is 81.13%, the precision is 81.37%, and the F1 is 81.02%. On the IMDB dataset, [13]suggested a stochastic gradient descent optimization approach with a linear function. As a consequence, the performance accuracy is 80.4%, the recall is 72.45%, the precision is 86.21%, and the F1 is 78.74%. On the massachusetts highways dataset, [14]suggested a stochastic gradient descent optimization technique with a tangent hyperbolic function. As a consequence, the performance accuracy is 98%, the recall is 82%, the precision is 93%, and the F1 is 87%.

On the WDBC dataset, [15]suggested an adaptive moment estimation optimization technique with a corrected linear unit function. As a consequence, the performance accuracy is 90.64%, the recall is 91%, the precision is 91%, and the F1 is 90%. On the WBCD dataset, the authors[16] suggested an adaptive moment estimation optimization technique with a sigmoid function. As a consequence, the performance accuracy is 97.2%, the recall is 97%, the precision is 97.5%, and the F1 is 97%. On the diagnostic breast cancer dataset, [17] developed an adaptive moment estimation optimization technique using a linear function. As a consequence, the performance accuracy is 96.5%, the recall is 98.9%, and the precision is 98.9%. On the NSL-KDD dataset, [18]suggested an adaptive moment estimation optimization technique utilizing a tangent hyperbolic function. As a consequence, the performance accuracy is 82.95%, the recall is 92.19%, the precision is 79.19%, and the F1 is 80.43%.

On the LiDAR dataset, [19]suggested a limited-BFGS-B optimization technique using a corrected linear unit function. As a consequence, the performance accuracy is 96.1%, the recall is 96.1%, the precision is 96.1%, and the F1 is 96.1%. On the breast cancer Wisconsin diagnostic dataset, the authors[20] proposed a limited-BFGS-B optimization technique using a sigmoid function. As a consequence, the performance accuracy is 96.5%, the recall is 97.5%, and the precision is 96.9%. On the PBD dataset,[21]suggested a limited-BFGS-B optimization technique with a linear function. The recall is 83.9%, the precision is 83.1%, and the F1 is 83.4% as a consequence. [22]developed a limited-BFGS-B optimization technique for the KDDCU99 dataset. As a consequence, the performance recall is 77.15%, the accuracy is 75.38%, and the F1 is 76.25%.

3. Proposed Methodology

The detailed workflow for this method was developed using a neural network. In this method, hidden layers are first linked in a 4:10:10:4 fashion, and then activation functions are calculated using the rectified linear unit function. Stochastic gradient descent optimizer uses these active outputs to form opinions. Finally, all of the mechanisms are linked and presented as rectified linear unit with stochastic gradient descent model.

3.1. Preprocessing

- **Normalization:** It is a scaling method in which values are shifted and rescaled until the value becomes between 0 and 1. If we know how our data is distributed then this technique will become more utilized.
- **Standardization:** In this technique, the values fall near and around the mean. Standardization makes the attributes mean to zero and the remaining distribution has a value of one as standard deviation.
- **Selected gene**: We are choosing only high-quality genes and also eliminating genes where two genes are mixed and contain very less amount of information.
- **logarithm Scale:** To work with a wide range of quantities and to draw an analysis of data on a graph, the logarithm scale is the best suitable nonlinear scale. In this, every interval is increased by a factor of logarithm base normally used one is base ten.

3.2. Hidden layer

In neural networks, hidden layers exist between the input layer and output layer. It is the most complicated component of any neural network, to find how many neurons are there[23]. Finally, none of them received high accuracy ratings. However, because we considered the hidden layer in our model as 4:10:10:4, we obtained a high accuracy value.

3.3. Rectified linear unit

In our proposed model, we used the ReLu activation function because it activates all neurons at the same time, whereas others activate neurons one by one. Because ReLu returns zero for all negative inputs, likely, any given unit will not activate at all, resulting in a sparse network.

$$f(r) = \begin{cases} 0, & \text{if } r < 0 \\ r, & \text{otherwise} \end{cases} \tag{8}$$

3.4. Stochastic gradient descent

Stochastic Gradient Descent (SGD) was used as a solver in our model; it is a type of weighted optimization algorithm that is used to perform parameter updates. for every training $a^{(i)}$ and label $y^{(i)}$ is

$$\theta = \theta - \eta \cdot \nabla_\theta S(\theta; a^{(i)}; b^{(i)}). \tag{9}$$

3.5. Regularization

We used the regularisation technique, which is a set of techniques that prevents overfitting in neural networks and thus improves deep learning model performance when dealing with new data. We can reduce model errors by implementing regularisation techniques in neural networks. Overfitting is less common in our model. This leads to training, where the neural network performs well with fewer errors, resulting in high accuracy. It has L1 and L2 regularisation types, but in our model L2 concept is used [24]. The following equation is provided.

L2 regularizer:

$$Costfunction = Loss + \lambda + \Sigma||W|| \tag{10}$$

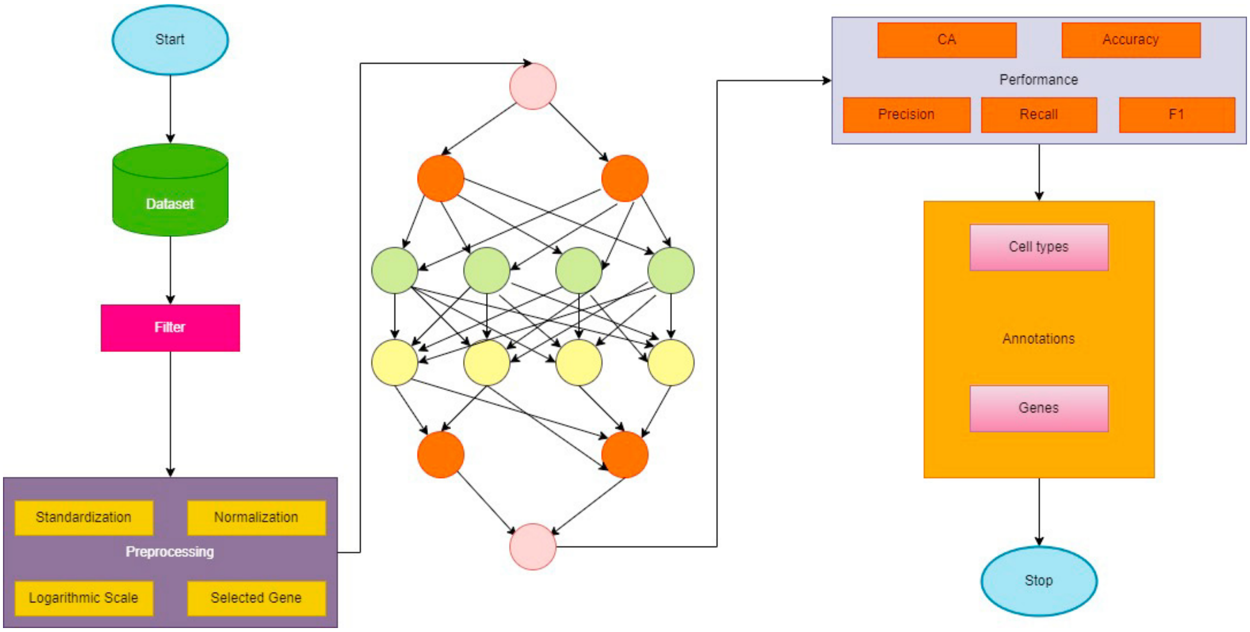


Fig. 1: Architectural diagram of proposed methodology.

4. Results and Discussion

We used single-cell RNA datasets which were created by Li Huipeng et al. (2017) and this dataset contains information about human colorectal tumors[25]. We can download the whole dataset from NCBI and the gene sequence number is GSE81861 other types of datasets can be used for implementing the same model. Another type of dataset's detailed information is given in table 1.

4.1. Filter

A filter is used to remove the genes and cells which is overlapping, dead, or infrequently by selecting genes in Fig.2.[26].

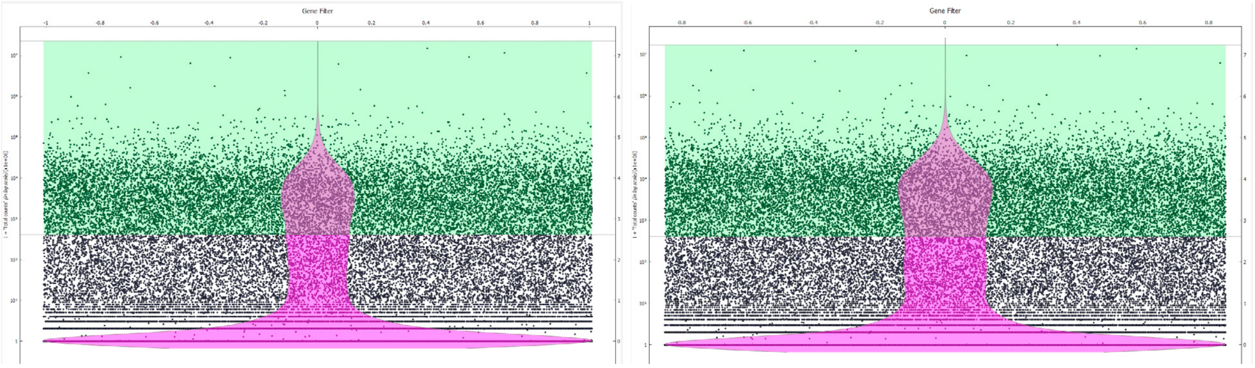


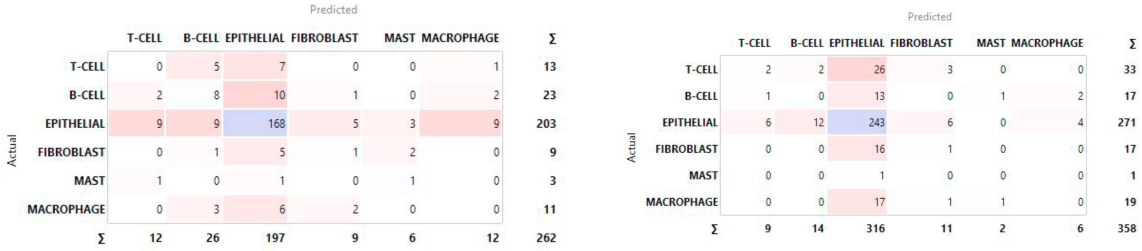
Fig. 2: Filter for (a) Normal mucosa; (b) Colorectal tumor

Table 1: List of Datasets and Comparision.

Dataset	Model (Activation function + Optimization algorithm)	Accuracy	F1	Precision	Recall
Human Colorectal Tumors	**ReLu + SGD**	**99.96%**	**98.40%**	**98.49%**	**98.54%**
	ReLu + Adam	99.75%	95.56%	95.50%	95.67%
	ReLu + Limited-BFGS-B	99.83%	96.76%	96.75%	96.79%
	Tanh + Adam	99.86%	97.14%	97.15%	97.17%
	TanH + SGD	99.86%	96.35%	96.38%	96.53%
	TanH + Limited-BFGS-B	99.84%	96.47%	96.48%	96.52%
	Sigmoid + Adam	99.86%	97.07%	97.09%	97.11%
	Sigmoid + SGD	99.74%	94.62%	93.97%	95.29%
	Sigmoid + Limited-BFGS-B	99.79%	96.47%	96.46%	96.48%
	Linear + Adam	99.76%	95.88%	95.96%	95.91%
	Linear + Limited-BFGS-B	99.55%	95.19%	95.21%	95.25%
	Linear + SGD	99.85%	96.76%	96.81%	96.81%
Hurricane Harvey	ReLu + SGD	95%	97%	97%	96%
Local Medical Center	Sigmoid + SGD	81.13%	81.37%	81.13%	81.02%
IMDB	Linear + SGD	80.4%	86.21%	72.45%	78.74%
Massachusetts Roads Dataset	TanH + SGD	98%	93%	82%	87%
WDBC	ReLu + Adam	90.64%	91%	91%	90%
WBCD	Sigmoid + Adam	97.2%	97.5%	97%	97%
Diagnostic Breast Cancer Dataset	Linear + Adam	96.5	98.9%	98.9%	N/A
NSL-KDD	TanH + Adam	82.95%	79.73%	92.19%	80.43%
LiDAR Data	ReLu + Limited-BFGS-B	96.1%	96.1%	96.1%	96.1%
Breast Cancer Wisconsin Diagnostic	Limited-BFGS-B + Sigmoid	96.5%	96.9%	97.5%	N/A
PDB	Linear + Limited-BFGS-B	N/A	83.9%	83.1%	83.4%
KDDCUP99	TanH + Limited-BFGS-B	N/A	75.38%	77.15%	76.25%

4.2. Performance measurement

- **Confusion matrix (CM):** It is a straightforward but extremely useful concept in deep learning. It is a matrix used to assess the efficiency of a classification model. Using any one of the deep learning models compare actual target values and predicted values in Fig.3.


Predicted
Actual
Predicted
Actual

	T-CELL	B-CELL	EPITHELIAL	FIBROBLAST	MAST	MACROPHAGE	Σ
T-CELL	0	5	7	0	0	1	13
B-CELL	2	8	10	1	0	2	23
EPITHELIAL	9	9	168	5	3	9	203
FIBROBLAST	0	1	5	1	2	0	9
MAST	1	0	1	0	1	0	3
MACROPHAGE	0	3	6	2	0	0	11
Σ	12	26	197	9	6	12	262

	T-CELL	B-CELL	EPITHELIAL	FIBROBLAST	MAST	MACROPHAGE	Σ
T-CELL	2	2	26	3	0	0	33
B-CELL	1	0	13	0	1	2	17
EPITHELIAL	6	12	243	6	0	4	271
FIBROBLAST	0	0	16	1	0	0	17
MAST	0	0	1	0	0	0	1
MACROPHAGE	0	0	17	1	1	0	19
Σ	9	14	316	11	2	6	358

Fig. 3: Confusion matrix for (a) Normal mucosa; (b) Colorectal tumor

As a result, we can determine whether or not our classification model is performing well.For instance, a 2X2 CM have elements like $True_{Positive}$, $True_{Negative}$, $False_{Positive}$, and $False_{Negative}$. We can calculate important performance measurement metrics like Accuracy, Precision or specificity, Recall or sensitivity, and F1 using the elements of the matrix. The formulas for these metrics are shown below.

- **Accuracy:** It is a metric used to explain to the model how it works across all the classes, and it gives the best results when all the classes are having the same priority. It is calculated as the proportion of correct predictions among all observations mentioned as in Table 1. The mathematical formula is shown below.

$$Accuracy = \frac{True_{Positive} + True_{Negative}}{True_{Positive} + True_{Negative} + False_{Positive} + False_{Negative}} \tag{11}$$

- **Precision**: It is the ratio of true positive cases and all cases where the prediction is correct. The precision of the model measures its accuracy in classifying a sample as positive and it mentioned in Table 1.

$$Precision = \frac{True_{Positive}}{True_{Positive} + False_{Positive}} \tag{12}$$

- **Recall:** It is the proportion of correctly identified positive cases. The recall of the model measures its ability to detect positive samples. The more positive samples detected, the higher the recall and it is shown in Table 1.

$$Recall = \frac{True_{Positive}}{True_{Positive} + False_{Negative}} \tag{13}$$

- **F1:** This is the balanced ratio between specificity or precision and sensitivity or recall, so it provides a synthesis of these two metrics. It is greatest when precision equals recall and mentioned in Table 1.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{14}$$

4.3. Annotation

Annotation defines organisms using cell types and their controlling regions, and each cell type contains information about the various type of genes.

- **Gene Markers**: The gene is a component of the double helix DNA molecule that carries the genetic information of a living organism but does not participate in the cell's molecular activities. Discover what makes genes unique in Table 2.
- **Cell types**: Annotating cell types is a critical step to define cell types to gene groups that are similar in Fig.4, and Table 2.

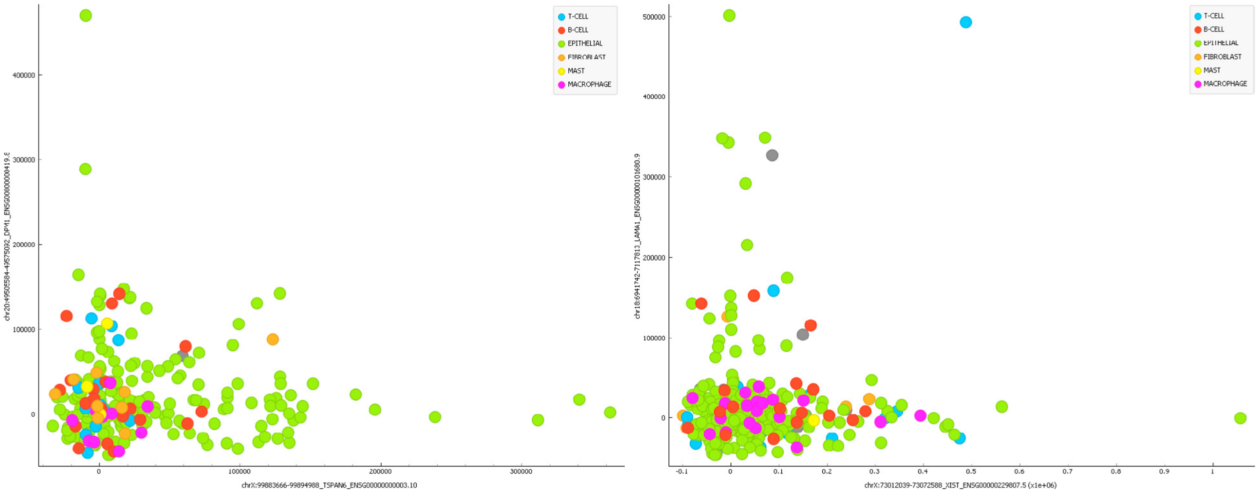


Fig. 4: Cell types for (a) Normal mucosa; (b) Colorectal tumor

Table 2: Cell types and gene markers for normal musoca and colorectal tumor.

Clusters	Cell types	Top five gene markers for normal musoca	Top five gene markers for colorectal tumor
0	T Cell	RPL15P3,TP53,TSSC2,DECR1,RP4-777D9.2	LAMA1,PER2,DENND4C,PPM1B,PSMB6
1	B Cell	RP11,PRSS27,RP11,POLR2G,RAN	CDKN1B,TAC01,ZNF713,PHC1P1,GPRIN2
2	EPITHELIAL	SAPCD1,TSPAN6,DDR1,PLEKHA1,SOX9	TGFBR3,FCRL2,SLC14A2,MMP7,KIF23
3	FIBROBLAST	KDELR3,NHP21,DPM1,LRP5L,RP11	PTBP2,FRMD8,CYPY4,KRT6B,AC092933.4
4	MAST	CD44,TCEAL1,COMMD3,ORMDL2,WDR83	PSMB6,RP11,XIST,RP11
5	MACROPHAGE	LAMP2,DHRS7,RPL4P4,ACO17099.3,PLP2	DNAJC17,MPV17,RNF6,APOE,CHP1

4.4. Heatmap

The heatmap displays the relationship between the individual cell types and genes that are clustered. And also heat map shows the expression matching between the cell types and genes Fig.5 [27].

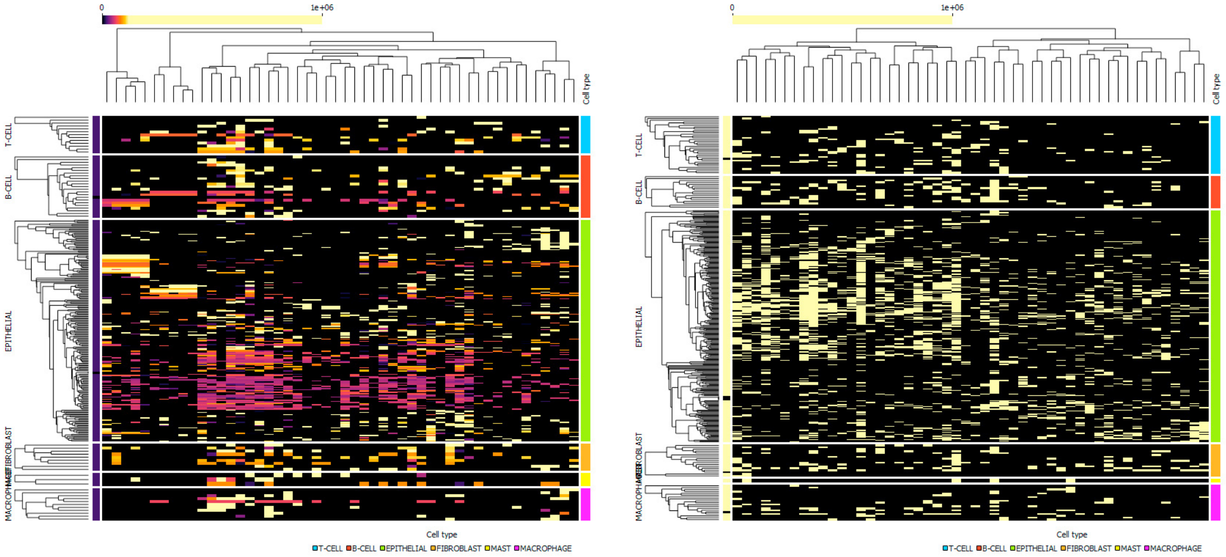


Fig. 5: Heatmap for (a) Normal mucosa; (b) Colorectal tumor

4.5. Violin plot(V-plot)

V-plot shows the data expression are assigned probability over the clusters. And also it will represent the gene expression level between each cell type in Fig.6 and Fig.7.

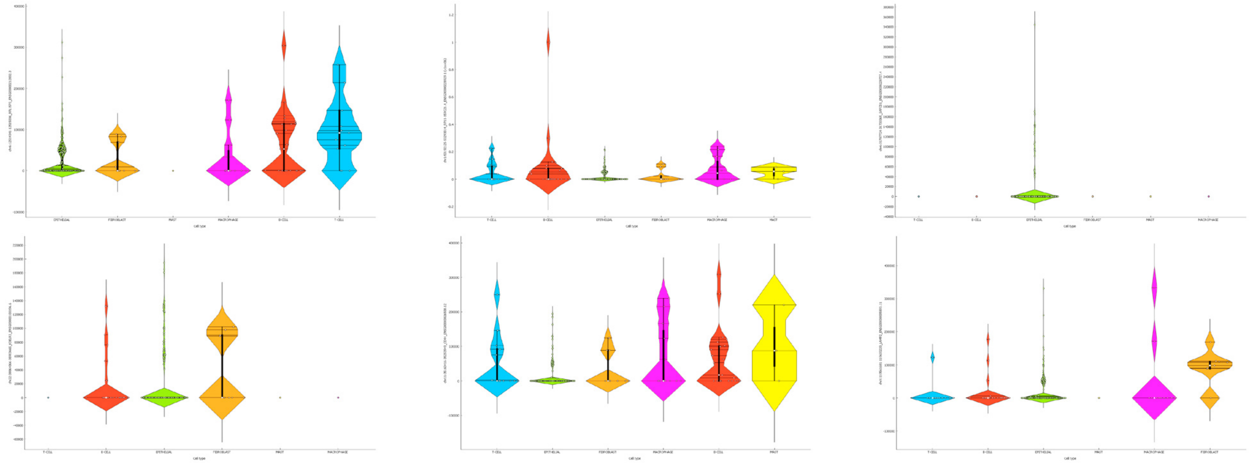


Fig. 6: V-plot for normal mucaso (a) RPL15P3 (T-cell), (b) RP11 (B-cell), (c) SAPCD1 (EPITHELIAL), (d) KDELR3 (FIBROBLAST), (e) CD44 (MAST) and (f) LAMP2 (MACROPHAGE)

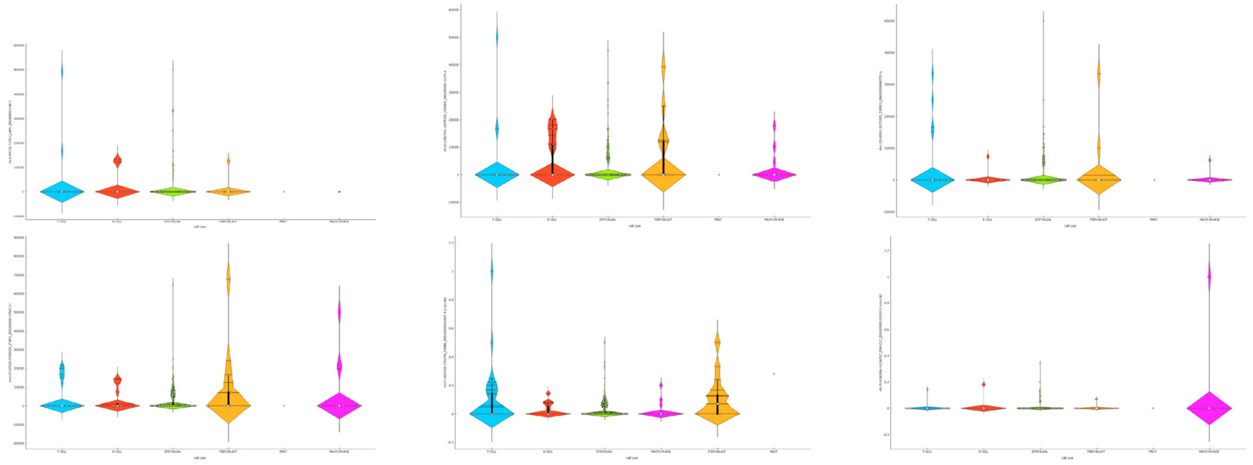


Fig. 7: V-plot for colorectal tumor (a) LAMA1 (T-cell), (b) CDKN1B (B-cell), (c) TGFBR3 (EPITHELIAL), (d) PTBP2 (FIBROBLAST), (e) PSMB6 (MAST) and (f) DNAJC17 (MACROPHAGE)

5. Conclusion & Future Work

In this paper, recently developed deep learning models were applied to single-cell RNA to analyse the unknown data, and significant understanding has been gathered.We believe that analysing transcriptome profiles is an important area in single-cell RNA dissection where deep learning will play a key role in the future. Our developed method produces more accurate results for determining each cell types and gene markers. This developed model will produce good clustering results, and this work demonstrates that identifying colorectal tumours and normal mucosa for single-cell RNA sequencing data using a rectified linear unit with stochastic gradient descent produces excellent results than other recently developed models. Other types of models, such as single or multi optimization can be designed in the future.

Acknowledgements

The work presented here falls under the Research Project Grant No. EEQ/2020/000104 and supported by the Department of Science and Technology (DST) and Science and Engineering Research Board (SERB), Government of India.

References

[1] Senapati, Biswajit and Ranjita Das.(2022) "Single Cell RNA Sequencing Data Analysis Using Machine Learning Techniques." *International Conference on Bio-Neuro Informatics models And Algorithm* **1** : 1–5.

[2] Dash, Rajashree, Rasmita Rautray, and Rasmita Dash (2020) "Utility of a Shuffled Differential Evolution algorithm in designing of a Pi-Sigma Neural Network based predictor model." *Applied Computing and Informatics* **16** (2): 1–19.

[3] Spears, and Brian K (2017) "Contemporary machine learning: a guide for practitioners in the physical sciences." *arXiv preprint arXiv:1712.08523* **1** (12): 1–29.

[4] Rasamoelina, Andrinandrasana David, Fouzia Adjailia, and Peter Sinčák. (2020) "A review of activation function for artificial neural network." *IEEE 18th World Symposium on Applied Machine Intelligence and Informatics* **7** (3): 281–286.

[5] Mourgias-Alexandris, George, A. Tsakyridis, N. Passalis, Anastasios Tefas, K. Vyrsokinos, and Nikolaos Pleros. (2019) "An all-optical neuron with sigmoid activation function." *Optics express* **27** (7): 9620–9630.

[6] Kalman, Barry L., and Stan C. Kwasny. (1992) "Why tanh: choosing a sigmoidal function." *Proceedings of IJCNN International Joint Conference on Neural Networks* **4**: 578–581.

[7] Senapati, Biswajit, and Ranjita Das.(2022) "Single Cell RNA Sequencing Data Analysing using Fuzzy De Based Clustering Technique." *Proceeding of International Conference on Advance In Distributed Computing and Machine Learning* **427**: 1–9.

[8] Newey, and Whitney K. (1988) "Adaptive estimation of regression models via moment restrictions." *Journal of Econometrics* **38** (3): 301–339.

[9] Gerber, Florian, and Reinhard Furrer. (2019) "optimParallel: An R package providing a parallel version of the L-BFGS-B optimization method." *R Journal* **11** (1): 352–358.

[10] Haddadpour, Farzin, Mohammad Mahdi Kamani, Mehrdad Mahdavi, and Viveck Cadambe. (2019) "Trading redundancy for communication: Speeding up distributed SGD for non-convex optimization." *International Conference on Machine Learning* **97**: 2545–2554.

[11] Kaur, Swapandeep, Sheifali Gupta, Swati Singh, Deepika Koundal, and Atef Zaguia.(2021) "Convolutional Neural Network based Hurricane Damage Detection using Satellite Images." *Soft Computing* **7** (2): 1–28.

[12] Zhang, Xiaoqing, Hongling Zhao, Shuo Zhang, and Runzhi Li. (2019) "A novel deep neural network model for multi-label chronic disease prediction." *Frontiers in genetics* **10** (2): 351–358.

[13] Teng, Zhaopu. (2021) "Teaching Machines How to Speak: Identifying Sensible Sentences from Phrases with Deep Neural Network." *The 5th International Conference on Compute and Data Analysis* **8** (1): 139–142.

[14] Zhang, Xiangrong, Xiao Han, Chen Li, Xu Tang, Huiyu Zhou, and Licheng Jiao. (2019) "Aerial image road extraction based on an improved generative adversarial network." *Remote Sensing* **11** (8): 930–949.

[15] Agarap, and Abien Fred. (2018) "Deep learning using rectified linear units (relu)." *arXiv preprint arXiv:1803.08375* **12** (5): 1–7.

[16] Gupta, Puja, and Shruti Garg. (2020) "Breast cancer prediction using varying parameters of machine learning models." *Procedia Computer Science* **171** (1): 593–601.

[17] Jais, Imran Khan Mohd, Amelia Ritahani Ismail, and Syed Qamrun Nisa. (2019) "Adam optimization algorithm for wide and deep neural network." *Knowledge Engineering and Data Science* **2** (1): 41–46.

[18] Masum, Mohammad, Hossain Shahriar, Hisham Haddad, Md Jobair Hossain Faruk, Maria Valero, Md Abdullah Khan, Mohammad A. Rahman, Muhaiminul I. Adnan, Alfredo Cuzzocrea, and Fan Wu. (2021) "Bayesian Hyperparameter Optimization for Deep Neural Network-Based Network Intrusion Detection." *IEEE International Conference on Big Data* **6** (3): 5413–5419.

[19] Kuck, Tahisa Neitzel, Edson Eyji Sano, Polyanna da Conceição Bispo, Elcio Hideiti Shiguemori, Paulo Fernando Ferreira Silva Filho, and Eraldo Aparecido Trondoli Matricardi.(2021) "A Comparative Assessment of Machine-Learning Techniques for Forest Degradation Caused by Selective Logging in an Amazon Region Using Multitemporal X-Band SAR Images." *Remote Sensing* **13** (17): 3341–3363.

[20] Kristoffersen Edward Mayce R., Lomboy, and Rowell M. Hernandez. (2021) "A comparative performance of breast cancer classification using hyper-parameterized machine learning models." *International Journal of Advanced Technology and Engineering Exploration* **8** (82): 1080–1101.

[21] Nallapareddy, Mohan Vamsi, and Rohit Dwivedula. (2021) "ABLE: Attention based learning for enzyme classification." *Computational Biology and Chemistry* **94** (5): 107558–17768.

[22] Yang, Chenming, Liang Zhou, Hui Wen, and Yue Wu. (2020) "U-ASG: A universal method to perform adversarial attack on autoencoder based network anomaly detection systems." *IEEE Conference on Computer Communications Workshops* **9** (3): 68–73.

[23] Huang, Guang-Bin, Yan-Qiu Chen, and Haroon A. (2000) "Classification ability of single hidden layer feedforward neural networks." *IEEE transactions on neural networks* **11** (3): 799–801.

[24] Enireddy, Vamsidhar, Mathe John Kenny Kumar, Babitha Donepudi, and C Karthikeyan. (2020) "Detection of COVID-19 using Hybrid ResNet and SVM." *In IOP Conference Series: Materials Science and Engineering* **993** (1): 012046–012056.

[25] (2017) "Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors." *Nature genetics* **49** (5): 708–718.

[26] Krishnaiah, Paruchuri R., and Laveen N. Kanal, eds. (1982) "Classification pattern recognition and reduction of dimensionality." *North-Holland Publishing Company* **2** (1): 1–903.

[27] Wang, Shuoguo, Constance Brett, Mohan Bolisetty, Ryan Golhar, Isaac Neuhaus, and Kandasamy Ravi (2019) "Single Cell Viewer (SCV): an interactive visualization data portal for single cell RNA sequence data." *BioRxiv* **10** (6): 664–789.