Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 9, 2025

SALIVA-BASED EARLY GASTRIC CANCER CLASSIFICATION USING DEEP SPARSE AUTOENCODERS

Muhammad Aqeel Aslam[1], Waheed Ur Rehman[*2], Faisal Bin Ubaid[3], Sana Liaquat[4], Aqsa Arshad[5], Ehtisham Lodhi[6]

[1]*School of Engineering & Applied Sciences, Electrical Engineering Department, Gift University, Gujranwala Pakistan,*
[*2]*School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China,*
[3]*Computer Science Department Sukkur IBA University, Sukkur, Pakistan*
[*2,4]*Department of Mechatronics Engineering, University of Chakwal, Chakwal 48800, Pakistan*
[5]*Department of Software Engineering, Riphah International University Islamabad*
[6]*Zhejiang University-University of Illinois Urbana-Champaign Institute, Haining, Zhejiang 314400, China*

[1]aqeel.aslam@gift.edu.pk,[*2]wrehman87@buaa.edu.cn, waheed.urrehman@uoc.edu.pk,
[3]faisalbinubaid@iba-suk.edu.pk, [4]Sana.liaquat@uoc.edu.pk, [5]67458@students.riphah.edu.pk,
[6]ehtishamlodhi@intl.zju.edu.cn

DOI:https://doi.org/10.5281/zenodo.17205221

Keywords
Deep Sparse Autoencoder; Gastric Cancer Classification; Saliva Samples; SoftMax Regression

Article History
Received: 04 July 2025
Accepted: 14 September 2025
Published: 26 September 2025

Corresponding Author: *
Waheed Ur Rehman

Abstract
Gastric cancer (GC) continues to represent a significant global health challenge, being one of the primary contributors to cancer-associated mortality. Its poor prognosis is largely attributed to the absence of clear early-stage symptoms and the shortage of reliable biomarkers. Consequently, there is a critical demand for rapid, accurate, and non-invasive diagnostic methods that can support early detection, prognosis, and clinical decision-making. Saliva has recently emerged as a valuable biofluid due to its non-invasive collection and diverse molecular composition. In this work, we introduce a deep learning framework designed to classify individuals into three categories: patients with early gastric cancer (EGC), those with advanced gastric cancer (AGC), and healthy controls. Our dataset consists of 300 saliva samples, including 90 from EGC patients, 100 from AGC patients, and 110 from healthy volunteers. The findings demonstrate that the proposed approach effectively differentiates EGC and AGC cases from healthy individuals, highlighting its strong potential for clinical translation. The FC-DSAE + SMC neural network achieved outstanding performance, yielding an accuracy of 99%, selectivity of 98.9%, sensitivity of 100%, specificity of 99.57%, detection rate of 99.76%, and an F-measure of 99.5% for the EGC category.

INTRODUCTION

Gastric cancer (GC) is considered a highly aggressive malignancy, ranking as the fourth most prevalent cancer worldwide and the second leading contributor to cancer-related mortality. Its early-stage detection remains particularly challenging [1–3]. GC is closely associated with genetic, lifestyles, and the environment [4]. According to the previous studies, northeast Asia has the highest incidents of GC. China has been severely affected by GC, where the number of GC cases accounts for 42% of the worldwide cases [5]. Generally, GC is divided into two stages, which are the early gastric cancer stage (EGC) and the advanced gastric cancer stage (AGC). The survival rate is closely

related to cancer stages [6]. If the cancer is diagnosed in the early stages, the mortality rate can be reduced to a great extent. The five-year survival rate is nearly 83% for EGC patients, whereas it is as low as 24% for AGC patients [7]. Only Japan has a reasonable survival rate for EGC patients which is more than 90%, this can achieve through early screening of cancer [8]. Therefore, achieving an early prognosis is crucial for enhancing patient survival rates. Although several new clinical techniques for cancer screening have been introduced in recent years [9–14], there remains a pressing need for approaches that can identify early gastric cancer (EGC) cases through methods that are non-invasive, user-friendly, rapid, and potentially suitable for self-testing at home. Saliva is a complex mixture that not only contains proteins, but also VOCs, DNA, RNA, fatty acids, and microorganisms, including important biomarkers associated with diseases, can reflect real-time human health status. Saliva diagnosis is a noninvasive, simple, fast, safe, and home self-diagnosis method that owns clinical translational prospect [15]. In our earlier study, specific amino acid biomarkers identified in human saliva were utilized to differentiate gastric cancer patients from healthy individuals [16]. Statistical evaluation was conducted using the Mann–Whitney U test to compare metabolic variations between the two groups, and a diagnostic model was subsequently developed through logistic regression. This approach demonstrated promising performance, yielding a specificity of 87.7% and a sensitivity of 80% [16]. Nevertheless, improving both sensitivity and specificity for gastric cancer detection remains a significant challenge. To address this limitation, we explored the application of deep learning methods to establish a model capable of accurately and efficiently classifying GC patients. For biomarker detection, we employed a Surface-Enhanced Raman Scattering (SERS) sensor [17–18]. However, Raman signal analysis is often complicated by spectral noise, including interference from background fluorescence, sample contaminants, and cosmic rays, which can distort the baseline [19–20]. Therefore, appropriate preprocessing techniques are required to mitigate these issues. Artificial intelligence (AI) has increasingly been integrated into the field of medical diagnostics, with machine learning (ML) approaches reported to enhance cancer detection accuracy by approximately 15–20% [21]. The remarkable progress in this area is largely attributed to the development of deep neural networks (DNNs) [22]. DNNs consist of multiple layers of interconnected neural units and include models Including architectures such as multilayer perceptrons (MLPs), convolutional neural networks (CNNs), restricted Boltzmann machines (RBMs), and autoencoders (AEs). MLPs are typically employed for classification tasks [23], while CNNs, inspired by the structure of the human visual cortex, are primarily used in image recognition [24]. RBMs, on the other hand, are probabilistic models designed to capture the distribution of input data [25]. Autoencoders, which exist in various forms, are widely applied to generate compact representations of data [26]. Their applications span a broad range, including latent feature extraction from text, image quality enhancement, dimensionality reduction, tomographic reconstruction in positron emission tomography, and hash function–based data mining [27–31].

In recent years, computational approaches have been increasingly employed to support pathologists in the interpretation of microscopic images [32–33]. Deep neural networks, in particular, have exhibited outstanding performance in the domains of pattern recognition and computer vision, achieving cutting-edge accuracy in image classification and recognition [34]. For example, Oikawa et al. [35] developed a multistage gastric cancer detection system using pathological images, which achieved a false-positive rate of 14.1%. Similarly, Xu et al. [36] applied convolutional neural network (CNN)–based segmentation and classification to histopathological images by separating them into epithelial and stromal regions, where CNNs were used to identify specific image patches. Wang et al. [37] proposed a CNN-based method for breast cancer detection that followed a comparable patch-based approach. Li et al. [38] also designed a patch-based deep learning framework for gastric cancer diagnosis, incorporating both shallow and deep networks for

classification. Beyond image-based studies, Daniel et al. [39] introduced a technique that classified gastric cancer using breath analysis, employing single- and multi-layer networks with the Resilient algorithm to achieve 92.54% accuracy. In another direction, Guyon et al. [40] presented a gene-selection-based approach for cancer detection, applying support vector machines with recursive feature elimination (RFE) and reporting an accuracy of 86% using the baseline model. More recently, Aslam et al. [41] give an idea of deep learning framework that utilized breath samples to distinguish early gastric cancer from advanced gastric cancer, and healthy individuals, achieving an overall accuracy of 97.3%.

In this work, to extract communal features from the Raman based salivary samples, we propose here a deep neural network comprises of Deep Sparse Autoencoder (DSAE) and SoftMax Classifier (SMC). DSAE constructed a Deep Neural Network (DNN), which was used as a classifier to discriminate between cancer-diagnosed patients and healthy subjects. In the proposed deep neural network, DSAE reconstructed the input by learning hierarchal features from the unlabeled Raman data. Once the deep sparse autoencoder has learned the features, a layer of these autoencoders is connected to an additional classifier. In this way, a cascaded deep neural network was formed. Lastly, we used this trained deep neural network as a classifier to predict the class labels of the input data. The proposed system for gastric cancer classification yields an accuracy of 99%. Moreover, the proposed FC-DSAE+SMC based neural network has distinguished 98%, 98.9% and 100% AGC, EGC and Healthy person respectively.

2. MATERIALS AND METHODS

Deep Sparse Autoencoder (DSAE): Machine learning (ML) techniques are broadly categorized into supervised and unsupervised approaches. Autoencoders fall under the unsupervised learning paradigm, as they are designed to capture correlations within high-dimensional input data to generate more effective feature representations. An autoencoder is an asymmetric, multilayer feed-forward neural network trained through backpropagation to reproduce its input at the output layer. As illustrated in Figure 1,

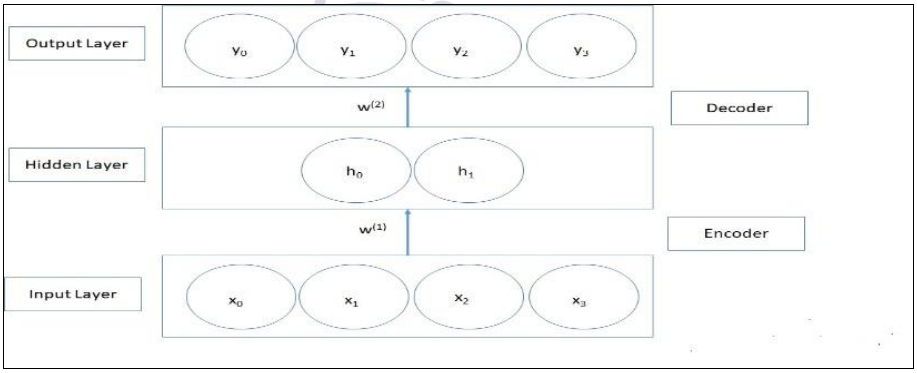

Figure 1. Basic structure of an autoencoder showing the encoder, hidden layer (h), and decoder, where input (x) is reconstructed as output (y) with weights w(1) and w(2).

a basic autoencoder consists of input, hidden, and output, with neurons connected via synapses. Its structure is typically divided into two main components— The encoder functions by mapping the input into a latent representation, whereas the decoder utilizes this representation to reproduce the initial data. In the figure, x denotes the input, h the hidden layer representation, and y the reconstructed output. Extending this concept, by stacking successive sparse autoencoders, a deep autoencoder is obtained, where the output of each preceding layer is utilized as the input for the succeeding one. In this study, a fully connected deep sparse autoencoder (FC-DSAE) was developed to classify gastric cancer using salivary Raman spectra. The overall framework for feature-communal learning based on FC-DSAE combined

with a SoftMax Classifier (SMC) is depicted in Figure 2

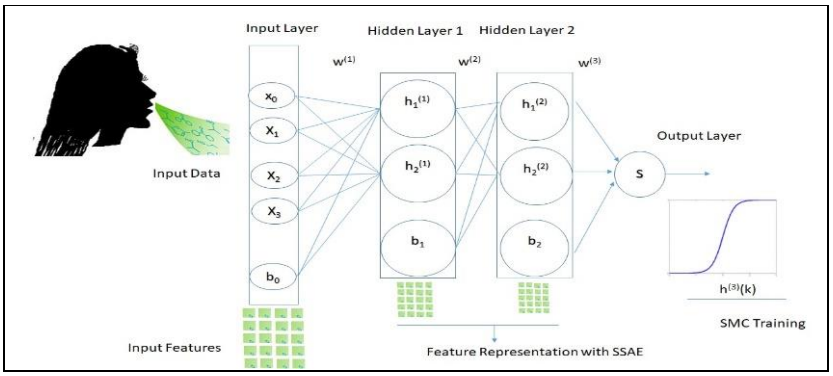

Figure2. Gastric cancer classification using FC-DSAE + SMC neural network from saliva samples

SoftMax Classifier (SMC):

The SoftMax classifier belongs to the supervised learning algorithm, which is represented as the following equation.

$$f_W = 1/(1 + \exp(-xW^T))$$

Where $f(.)$ is the sigmoid function, SMC is used to classify the data of multiple classes. SMC is the advanced and generalized version of logistic regression (42).

Clinical Sample Verification: In conducting this study, compliance with the Reporting Recommendations for Tumor Marker Prognostic Studies guidelines was ensured. A total of 220 saliva samples were analyzed using Surface-Enhanced Raman Spectroscopy (SERS) technology. The dataset, obtained from Horiba Ltd., Shanghai, China, was categorized into three groups: patients with early gastric cancer (EGC), patients with advanced gastric cancer (AGC), and healthy individuals. For model development, the dataset was split into training and testing subsets, with 70% allocated for training and 30% reserved for evaluation. The detailed distribution of samples across the three groups is presented in Table I.

Table I. The clinical characteristics of volunteers

Group	Number of Volunteers	Age (Years)	Gender (M:F)
AGC	100	60 ± 9	61:39
EGC	90	53 ± 8.6	56:34
Healthy	110	35 ± 10	57:43

Feature Communal Learning based on Deep Sparse Autoencoder + SoftMax Classifier (FC-DSAE+SMC): This model comprises of Feature Communal Learning based on Deep Sparse Autoencoder with SoftMax Classifier. The DSAE was initially assigned a set of hyper-parameters. These parameters include the network parameters, the weight of SMC, and the regularization term. For evaluation, the proposed model is tested on the validation dataset, and the classification outcomes are used to optimize the hyperparameters. The training framework is organized into three distinct steps

In the first step, SAE was employed to the Raman data of the saliva sample in training set X, to learn the basic representations of the h(1)(x) on the data by updating all the weights w(1).

These primary representations are then fed to the sparse autoencoder to find the secondary representation h(2) by adjusting the weights w(2).

These secondary representations now act as the input to the SMC, and it was trained for mapping the h(2) by adjusting the weights w(3).

These three layers were integrated to construct the FC-DSAE+SMC model, which demonstrated the

capability to classify cancerous from non-cancerous data. The FC-DSAE component was trained in a bottom-up manner using an unsupervised approach, followed by the SMC module, which applied a supervised learning strategy to train the top layer and fine-tune the overall architecture.

Ethical Approval: The study was approved by the Ethical Committee of Shanghai Jiao Tong University (Shanghai, China), and informed consent was obtained from all patients in line with clinical research guidelines.

3. Experimental Results and Discussion

Performance Evaluation: The effectiveness of the proposed Feature Communal Learning model based on DSAE + SMC was assessed by comparing its performance with existing methods using multiple evaluation metrics. The assessment was conducted through standard quantitative parameters, including sensitivity, specificity, selectivity, detection rate, and F-measure. In addition, performance was measured in terms of precision (Pr), recall or true positive rate (TPR), false positive rate (FPR), and F-measure.

For these calculations, True Positive (TP) and True Negative (TN) represent the number of cases correctly classified as malignant and benign, respectively. Conversely, False Positive (FP) refers to benign cases incorrectly identified as malignant, while False Negative (FN) corresponds to malignant cases misclassified as benign. The formulas used to compute these evaluation metrics are given as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} * 100\ \%$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} * 100\ \%$$

$$\text{Selectivity} = \frac{\text{TP}}{\text{TP} + \text{FP}} * 100\ \%$$

$$\text{Detection rate} = \frac{\text{Sensitivity} + \text{Specificity}}{2} *\ 100\ \%$$

$$\text{F} - \text{measure} = 2\left(* \frac{\text{Selectivity} * \text{Sensitivity}}{\text{Selectivity} + \text{Sensitivity}}\right) *\ 100\%$$

Computational Considerations: In this study, we have used a CPU Intel (R) Core (TM) i3, with 2.3GHz processor and 8GB of RAM for the implementation of all the stated machine learning algorithms. The implementation of all the algorithms of machine learning was performed on MATLAB 2017a, LabSpec. The training+ validation data set consists of 255 samples, and test dataset consist of 45 samples. The test set included 10 samples of EGC, 13 samples of AGC and 22 samples of healthy persons.

Artificial Neural Network (ANN): We also developed an ANN using scaled conjugate backpropagation technique. This method also used to analyze the performance of the systems. We trained the model; after completing the training we tested the developed model on our dataset. Again, we keep the testing dataset to 15% of the total dataset, whereas, the remining dataset was used for training and validation.

ISSN (e) 3007-3138 (p) 3007-312X

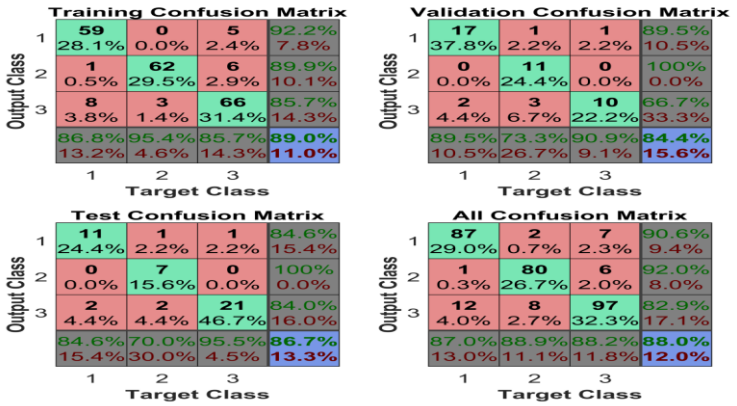

Figure 3 Confusion Matrix for ANN based Classification of AGC, EGC and Healthy Persons

Figure 3 shows the results in form of confusion matrix and receiver operating characteristics for the classification of EGC, AGC and healthy persons using ANN based system. In this technique, the model has distinguished 87, 80 and 97 samples correctly as AGC, EGC and Heatly persons. The training dataset has an accuracy of 89%, validation dataset has an accuracy of 84.4% and test dataset has an accuracy of 86.7%. The ANN based neural network produced an accuracy of 88%, selectivity 88.8%, sensitivity 91.95%, specificity 95.19%, detection rate 93.57% and F-measure 90.34%. All these parameters were calculated for EGC class.

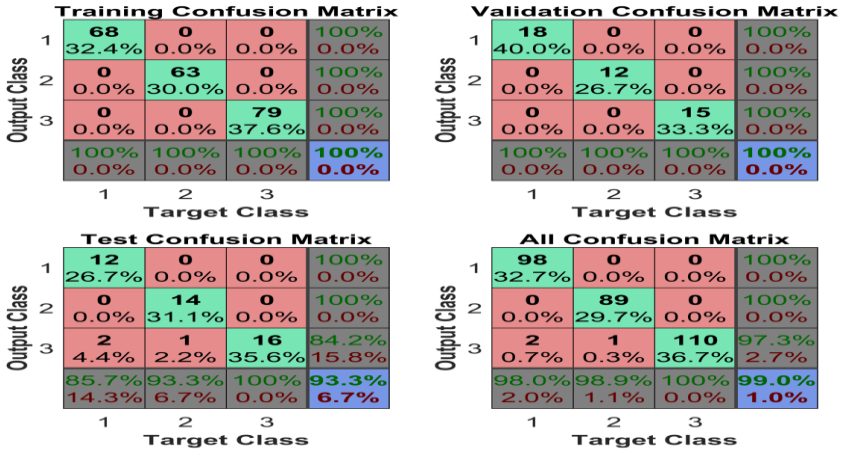

Figure 4 FC-DSAE + SMC based Classifier System for AGC, EGC and Healthy Persons

Figure 4 show the result of the developed FC-DSAE + SMC based classifier, which distinguishes the AGC, EGC and Healthy samples more accurately. This model produces an overall accuracy of 99%on the complete dataset. This model misclassified only 3 samples in the test data out of 45 samples. This model produced an accuracy of 93.3% on the test dataset. In the training phase, the FC-DSAE + SMC has distinguished all the AGC, EGC and Healthy samples accurately producing an accuracy of 100%. This model has misclassified 2 samples in AGC and 1 sample from the EGC class. Moreover, this method has distinguished all the healthy samples accurately. The FC-DSAE + SMC based neural network produced an accuracy of 99%, selectivity 98.9%, sensitivity 100%, specificity 99.57%, detection rate 99.76% and F-measure 99.5%. All these parameters were calculated for EGC class.

Table II. Comparison of FC-DSAE+SMC with ANN, in terms of Accuracy, Selectivity, Specificity, Sensitivity, Detection Rate, F-Measure

Method	Accuracy	Selectivity	Specificity	Sensitivity	Detection Rate	F-Measure
ANN	88.0%	88.8%	95.19%	91.95%	93.57%	90.34%
FC-DSAE+SMC Classifier	99%	98.9%	99.52%	100%	99.76%	99.5%

Table II summarizes the result of this study. This table indicates that our developed model is better in predicting the GC and non-GC samples with very high efficiency.The FC-DSAE+SMC based neural network has the highest accuracy, selectivity, sensitivity, specificity detection rate, and F-measures when compare to all state of art strategies shown in Table II. However, ANN based neural network has the low accuracy, selectivity, specificity, detection rate, and F-measures. Therefore, we can say that our proposed model using deep neural network architecture has achieved higher accuracy, higher precision, higher recall for the gastric cancer classification, and the method has outperformed all other existing techniques.

4. Conclusions

In this study, we have developed a deep neural network model using sparse auto encoder architecture. This model was used to detect gastric cancer by using saliva samples. The proposed model used the unsupervised technique to capture the high-level feature representation of the Raman data. The classifier works very well with very high accuracy by utilizing these high-level feature representations. From the qualitative analysis and quantitative analysis, we can conclude that the deep sparse autoencoder with the combination of SoftMax classifier has outperformed the other state of art methods in terms of efficiency, accuracy, selectivity, sensitivity, and classification which have been used in this study. The proposed model has accurately distinguished the gastric cancer patients from healthy persons.

References

Torre, Lindsey A., et al. Global cancer statistics, 2012. CA: a cancer journal for clinicians. 2015; 65.2: 87-108.

A Sheikh, Ishfaq, et al. A proteomics based approach for the identification of gastric cancer related markers. Current Pharmaceutical Design. 2016; 22.7: 804-811.

Qiu, Fu-Ming, et al. Mining novel biomarkers for prognosis of gastric cancer with serum proteomics. Journal of Experimental & Clinical Cancer Research. 2009; 28.1: 1-7.

Wadhwa, Roopma, et al. Gastric cancer—molecular and clinical dimensions. Nature Reviews Clinical oncology. 2013; 10.11: 643-655.

Washington, Kay. 7th Edition of the AJCC cancer staging manual: stomach. Annals of Surgical Oncology. 2010; 17.12: 3077-3079.

Ooki, Akira, et al. Phosphatase of regenerating liver-3 as a prognostic biomarker in histologically node-negative gastric cancer. Oncology Reports. 2009; 21.6: 1467-1475.

Alberts, Steven Robert, A. Cervantes, and C. J. H. Van de Velde. Gastric cancer: epidemiology, pathology and treatment. Annals of Oncology. 2003; 14 suppl. 2: ii31-ii36.

Asaka, Masahiro, and Katsuhiro Mabe. Strategies for eliminating death from gastric cancer in Japan. Proceedings of the Japan Academy, Series B. 2014; 90.7: 251-258.

Zhou, Chengmao, et al. A machine learning-based predictor for the identification of the recurrence of patients with gastric cancer after operation.Scientific Reports. 2021; 11.1: 1-7.

Sharma, Harshita, et al. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. Computerized Medical Imaging and Graphics. 2017; 61: 2-13.

Aslam, Muhammad Aqeel, et al. SVM based classification and prediction system for gastric cancer using dominant features of saliva. Nano Biomed Eng. 2020; 12.1: 1-13.

Taninaga, Junichi, et al. Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive check- data

A case-control study. Scientific Reports. 2019; 9.1: 1-9.

Chen, Yunsheng, et al. Breath analysis based on surface-enhanced Raman scattering sensors distinguishes early and advanced gastric cancer patients from healthy persons. ACS Nano. 2016; 10.9: 8169-8179.

Chen, Yiping, et al. One-step detection of pathogens and viruses: combining magnetic relaxation switching and magnetic separation. ACS Nano. 2015; 9.3: 3184-3191.

Hofman, Lindsay F. Human saliva as a diagnostic specimen. The Journal of Nutrition. 2001; 131.5: 1621S-1625S.

Chen, Yunsheng, et al. Salivary analysis based on surface enhanced Raman scattering sensors distinguishes early and advanced gastric cancer patients from healthy persons. Journal of Biomedical Nanotechnology. 2018; 14.10: 1773-1784.

Liu, Jinchao, et al. Deep convolutional neural networks for Raman spectrum recognition: a unified solution. Analyst. 2017. 142.21: 4067-4074.

Butler, Holly J., et al. Using Raman spectroscopy to characterize biological materials. Nature Protocols.2016; 11: 664-687.

Barman, Ishan, et al. Effect of photobleaching on calibration model development in biological Raman spectroscopy. Journal of Biomedical Optics. 2011; 16.1.

Zhang, Lin, and Mark J. Henson. A practical algorithm to remove cosmic spikes in Raman imaging data for pharmaceutical applications. Applied Spectroscopy. 2007; 61.9: 1015-1020.

Cruz, Joseph A., and David S. Wishart. Applications of machine learning in cancer prediction and prognosis. Cancer Informatics. 2006; 2: 59-78.

Hinton, Geoffrey E., Simon Osindero, and Yee-WhyeTeh. A fast-learning algorithm for deep belief nets. Neural Computation/ 2006; 18.7: 1527-1554.

D. W. Ruck, S. K. Rogers, M. Kabrisky, M. E. Oxley and B. W. Suter. The multilayer perceptron as an approximation to a Bayes optimal discriminant function. IEEE Transactions on Neural Networks. 1990; 1.4: 296-298. doi: 10.1109/72.80266.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25 (2012) (NIPS'2012).

Fischer, Asja, and Christian Igel. Training restricted Boltzmann machines: An introduction. Pattern Recognition. 2014; 47.1: 25-39.

Droniou, Alain, and Olivier Sigaud. Gated autoencoders with tied input weights.Proceedings of the 30th International Conference on Machine Learning, 2013, pp:154-162.

Wei C, Luo S, Ma X, Ren H, Zhang J, et al. Locally Embedding Autoencoders: A Semi-Supervised Manifold Learning Approach of Document Representation. PLOS ONE. 2016; 11.1: 1-20. e0146672. https://doi.org/10.1371/journal.pone.0146672

Lore, Kin Gwn, AdedotunAkintayo, and Soumik Sarkar. LLNet: A deep autoencoder approach to natural low-light image enhancement. Pattern Recognition. 2017; 61: 650-662.

Wang, Yasi, Hongxun Yao, and Sicheng Zhao. Auto-encoder based dimensionality reduction.Neurocomputing. 2016; 184: 232-242.

Cui J, Liu X, Wang Y, Liu H. Deep reconstruction model for dynamic PET images. PLOS ONE. 2017; 12.9: 1-21. e0184667. https://doi.org/10.1371/journal.pone.0184667

M. Á. Carreira-Perpiñán and R. Raziperchikolaei, Hashing with binary autoencoders. Proc. IEEE Conf. Comput. Vis. Pattern Recognit.,Boston, MA, USA, 2015, pp. 557-566.

Ghaznavi, Farzad, et al. Digital imaging in pathology: whole-slide imaging and beyond. Annual Review of Pathology: Mechanisms of Disease. 2013; 8: 331-359.

Gurcan, M. N., et al. Histopathological image analysis: a review. IEEE Rev Biomed Eng. 2009; 2: 147-171.

Russakovsky, Olga, et al. Imagenet large scale visual recognition challenge. International journal of Computer Vision. 2015; 115.3: 211-252.

K. Oikawa, A. Saito, T. Kiyuna, H. P. Graf, E. Cosatto and M. Kuroda, Pathological Diagnosis of Gastric Cancers with a Novel Computerized Analysis System. *Journal of Pathology Informatics. 2017; 8.2.*

Xu, Jun, et al. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomputing. 2016; 191: 214-223.

Wang, Dayong, et al. Deep learning for identifying metastatic breast cancer. (Jun, 2016) [Online]. Available: https://arxiv.org/abs/1606.05718

Y. Li, X. Li, X. Xie and L. Shen, Deep learning based gastric cancer identification, 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 2018, pp. 182-185, doi: 10.1109/ISBI.2018.8363550.

Daniel, D. Arul Pon, and K. Thangavel. Breathomics for gastric cancer classification using back-propagation neural network. Journal of Medical Signals and Sensors. 2016; 6.3: 172-182.

Guyon, I., Weston, J., Barnhill, S. &Vapnik, V. Gene selection for cancer classification using support vector machines.Mach. Learn. 2002;46. (1–3): 389–422.

Aslam, Muhammad Aqeel, et al. Breath analysis based early gastric cancer classification from deep stacked sparse autoencoder neural network. Scientific Reports, 2021; 11.1: 1-12.

Zhang, Yan, Erhu Zhang, and Wanjun Chen. Deep neural network for halftone image classification based on sparse auto-encoder. Engineering Applications of Artificial Intelligence. 2016; 50: 245-255.