

# Diagnosing Epidermal basal Squamous Cell Carcinoma in High-resolution, and Poorly Labeled Histopathological Imaging

**Mani Manavalan**

Technology Architect, Cognizant Technology Solutions, Teaneck, New Jersey, USA

\*Corresponding Contact:

Email: [manimanavalan47@gmail.com](mailto:manimanavalan47@gmail.com)

## ABSTRACT

The most appropriate method to uncover patterns from clinical records for each patient record is to create a bag with a variety of examples in the form of symptoms. The goal of medical diagnosis is to find useful ones first, and then map them to one or more diseases. Patients are often represented as vectors in some aspect. Pathologists and dermatopathologists diagnose basal cell carcinomas (BCC), one of the most frequent cutaneous cancers in humans, on a regular basis. Improving histological diagnosis by producing diagnosis ideas, i.e. computer-assisted diagnoses, is a hotly debated research topic aimed at improving safety, quality, and efficiency. Due to their improved performance, machine learning approaches are rapidly being used. Typical images obtained by scanning histological sections, on the other hand, frequently have a resolution insufficient for today's state-of-the-art neural networks. Furthermore, weak labels hamper network training because just a small portion of the image signals the disease class, while the majority of the image is strikingly comparable to the non-disease class. The goal of this work is to see if attention-based deep learning models can detect basal cell carcinomas in histological sections and overcome the ultra-high resolution and poor labeling of full slide images. With an AUC of 0.99, we show that attention-based models can achieve nearly flawless classification performance.

### Key words:

Basal cell carcinomas (BCC), Machine learning, Convolutional neural networks (CNN), Histopathological imaging

12/31/2020

Source of Support: None, No Conflict of Interest: Declared

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Attribution-NonCommercial (CC BY-NC) license lets others remix, tweak, and build upon work non-commercially, and although the new works must also acknowledge & be non-commercial.



## INTRODUCTION

Rather than controlled experiment conditions, data is collected from a variety of real-world applications in many real-world observational research. Data quality problems for this observational data include (i) data correctness, (ii) data completeness, (iii) data consistency, and (iv) data balance (Sun et al., 2018). More crucially, clinical judgments are made in real life based on only a few informative and valuable traits, i.e. features, rather than the complete patient record. The use of computer models to extract crucial information from enormous amounts of partial and low-quality data in order to create reliable diagnosis results has sparked widespread attention. Basal cell carcinomas (BCCs) are among the most frequent types of skin cancer in people (Chinem and Miot, 2011). BCCs are frequently diagnosed by pathologists and dermatopathologists due to their prevalence. In terms of safety, quality, and efficiency, digital pathology enhances and simplifies histological diagnosis (Griffin and Treanor, 2017). Pathologists' diagnoses are improved by digital pathology, which provides diagnostic support in the form of computer-assisted diagnostics (Komura and Ishikawa, 2018). Recently, machine learning technologies have largely been used to do this (Bynagari, 2019; Manavalan, 2019; Donepudi, 2014b; Ahmed and Neogy, 2009; Ahmed et al., 2011). Such technologies could help doctors, particularly pathologists, discover new histological patterns for diagnosing ailments.

Convolutional neural networks (CNN; LeCun et al., 1998; Krizhevsky et al., 2012) would be the established way to tackle this picture classification job from a machine learning perspective, owing to their success in analyzing biological and medical data. Examples include the diagnosis of melanoma on par with dermatologists (Esteva et al., 2017) or the prediction of cardiovascular risk factors using retinal fundus pictures (Esteva et al., 2017). The size of typical histological images obtained at a resolution sufficient to see cellular structures, on the other hand, is incompatible with existing state-of-the-art Convolutional neural networks designs. Typically, such images have a resolution of 50,000x100,000 pixels, whereas Convolutional neural networks are used on images with a maximum resolution of 4,096x4,096 pixels (Momeni et al., 2018). Recent attempts to train Convolutional neural networks on histopathologic pictures usually bypass the ultra-high resolution problem by sampling random image patches from the complete image (Albarqouni et al., 2016; Janowczyk and Madabhushi, 2016), which leaves the weak labeling problem unsolved. Komura and Ishikawa provide an overview of machine learning and deep learning strategies for tackling the high-resolution challenge of histopathology slides (2018).

### Problem Statement

Furthermore, histopathology slide classification suffers from a weak label problem: the entire picture slide is labeled with a single class (i.e. diagnosis), but a considerable portion of the image is identical in all classes, and only a tiny part is representative of the respective class. For assessing full slide images, multiple instance learning (MIL) and attention-based models have recently been developed (Ilse et al., 2018).

### Objectives of the Study

The aim of the project is to see if traditional machine learning approaches can be used to detect basal cell carcinomas in whole slide images (WSI) while overcoming the ultra-high resolution and weak label difficulties. Furthermore, we want to discover the picture key regions that are significant for the predictive model's decision, and then compare them to the diagnostic key regions for board-certified pathologists.

## LITERATURE REVIEW

### Multi-instance Learning

Weakly supervised learning (WSL) is the machine learning term for this situation, and multi-instance learning (MIL) is a common example (Zhou, 2017). Only the bag label is provided in this perspective, which treats the input sample as a bag of instances. MIL models enable for the prediction of new bag labels with containing instances through learning and training. MIL was first proposed by Dietterich et al. (1997) for drug molecule activity prediction and has since been widely used in a variety of fields, including medical imaging and video analysis (Ilse et al., 2018), syndrome differentiation in Traditional Chinese Medicine (TCM) (Wang et al., 2018), pulmonary embolism and colon cancer detection (Manivannam et al., 2017), and retinal nerve fiber layer visibility classification. MIL is used to diagnose a single ailment, i.e. binary classification on a single task.

The bag is labeled positive only if at least one instance is positive, according to the MIL definition; otherwise, the bag is labeled negative. Capturing correlations between examples and identifying the most informative instances are critical in this scenario. Many prior works, such as EM-DD (Zhang and Goldman, 2002), mi-SVM (Andrews et al., 2003), mi-Graph (Zhou et al., 2009), and miFV (Wei et al., 2014), focused on the latter, i.e., important instances detection. It's worth noting that in some cases, ignoring the relationships between instances can lead to erroneous predictions. The bag is designated as beach only if the sky, ocean, and sand are all present (Foulds and Frank, 2010). When dealing with a complicated process like a medical diagnosis, professionals must not only examine risk variables independently, but also consider the impact of their co-occurrence. In the model-building process, this is one of our starting places for measuring instance correlations (Bynagari & Fadziso, 2018; Manavalan, 2019; Donepudi, 2014a; Abedin et al., 2012; Ahmed & Dey, 2009a; Azad et al., 2011; Begum et al., 2012).

Ramon et al. (2000) were the first to propose a multi-instance neural network. This method computes instance probabilities, which are then processed by the log-sum-exp operator to obtain the bag probability, and the entire process is trained from beginning to end. Their work reveals how neural networks may tackle the MIL problem effectively and simply. More neural network-based MIL architectures have since been proposed for various applications (Zhang and Zhou, 2004; Yan et al., 2018; Bynagari, 2018; Manavalan, 2018; Donepudi, 2015; Ahmed, 2009; Neogy and Ahmed, 2015; Siddique & Ahmed, 2015; Ahmed & Dey, 2009b). In contrast to the method of calculating instance probabilities first, Wang et al. (2018) present a unique framework that derives the bag embedding using instance-level MIL pooling first, then builds a classifier based on it to calculate the bag probability. Their work introduces a novel method for creating multi-instance neural networks.

Furthermore, the attention method has demonstrated some performance gains in terms of capturing relationships between instances and between instances and bags. For the time being, it has been widely used in picture and text analysis (Xu et al., 2015; Lin et al., 2017; Manavalan, 2014; Bynagari & Amin, 2019; Amin & Manavalan, 2017; Fadziso & Manavalan, 2017; Donepudi, 2016) with two sub-categories: task-supervised attention and self-supervised attention. The first captures the relationships between the source and the target (Qi et al., 2017), while the second computes the source's intra-relationship (Vaswani et al., 2017). Both sub-categories are necessary for the MIL to function properly.

## MIL with Neural Networks

Because the MIL underlying function above is so flexible, we may simulate any transformation and score function as long as it is permutation-invariant. As a result, we use the neural network to parameterize a set of transformations. Let  $x$  be a bag of  $M$  instances, and the transformer  $\varphi_\tau$ , where  $\tau$  are parameters, turns instances into the embedding space with  $K$  dimensions, that is,  $V_{m,k} = \varphi_\tau(x_m), m \in M$ . The bag probability of  $x_m$  is thus calculated by the transformation:  $\theta_\omega: \eta_{\phi_k \in K}(v_{m,k}) \rightarrow [0, 1]$ . If you're using the bag-level MIL pooling method,  $\theta_\omega$  is an injective function, or else is parameterized by the neural networks using parameters  $\omega$ , and if you're using  $\phi$  the trainable MIL pooling methods, are also parameters.

Multiple Instance Learning aggregation Multiple Instance Learning pooling is the important step for bridging instances to bags, as demonstrated above, and different applications have varying preferences for Multiple Instance Learning pooling methods. The sole constraint they have on neural networks is that they must be differentiable. The instance-level multi-instance learning pooling options in MINet (Wang et al., 2018) are max pooling, mean pooling, and log-sum-exp pooling on each instance, while Yan et al (2018) developed a new dynamic pooling method integrating both instance-level and bag-level approaches.

## Attention-based Massive Multiple Instance Learning

By splitting complete slide images into patches of relatively tiny resolution that can be processed by typical Convolutional neural network architectures, we address the problem of classifying extremely high resolution images as a multiple instance learning problem. This raises the issue of credit assignment, or how to classify a whole WSI as 'contains Basal cell carcinomas or 'does not include Basal cell carcinomas by combining signals from a large number of patches. Simple methods, such as averaging patch predictions or adopting the maximally activated patch's forecast, have clear credit assignment issues. As a result, we use Ilse and colleagues' (Ilse et al., 2018) attention-based multi-instance learning pooling and compare it to a baseline of downscaled WSI as well as patch-based approaches employing mean and max pooling. Figure 1 depicts an overview of the patch-based technique. For all patch-based studies, we use the well-known VGG11 architecture by Simonyan and Zisserman (2014).

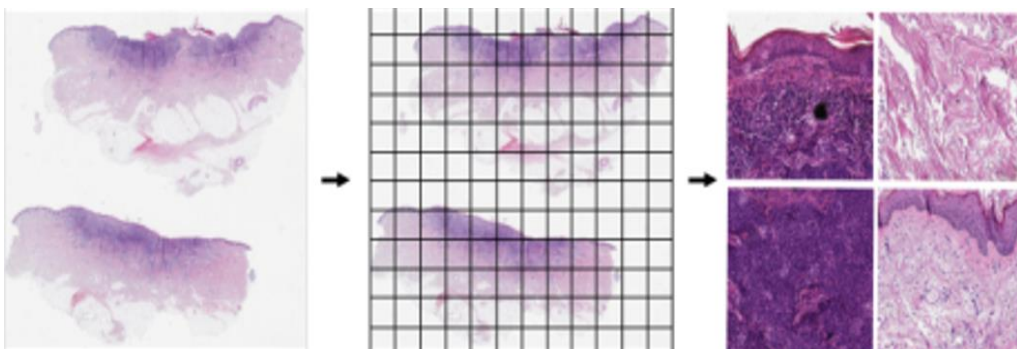


Figure 1: The input for the detection of BCC is histopathology slides with ultra-high resolution (typically  $>20,000 \times >20,000$  pixels). The complete image is divided into patches with a resolution of  $224 \times 224$  pixels in the center. Right: In a multiple instance learning setting, patches represent tiny sections of histopathology slides and are used as instances.

## Baseline method

A typical CNN trained on entire slide images down-scaled to 1024x1024 pixels is used as a baseline approach. The histological sections were classified using only 3% of the available data due to the down-sampling approach. Klombauer et al. (2017) developed a CNN that uses SeLU activation functions and comprises of five blocks of convolution-convolution-maxpooling. Manual hyperparameter tuning was used to optimize the architecture and hyperparameters of this CNN using a validation set. SGD was used to teach the network. Re-training the networks 100 times yielded the mean and standard deviation of accuracy, the F1 score, and the area under the ROC curve (AUC).

## Interpretation method

We used (1) Integrated Gradients (Sundararajan et al., 2017) as the baseline approach and (2) the learnt attention weights per image patch of the attention process mentioned above to interpret the neural network predictions. The attention weights were shown to establish the key regions for the CNN model's categorization (Figure 2).

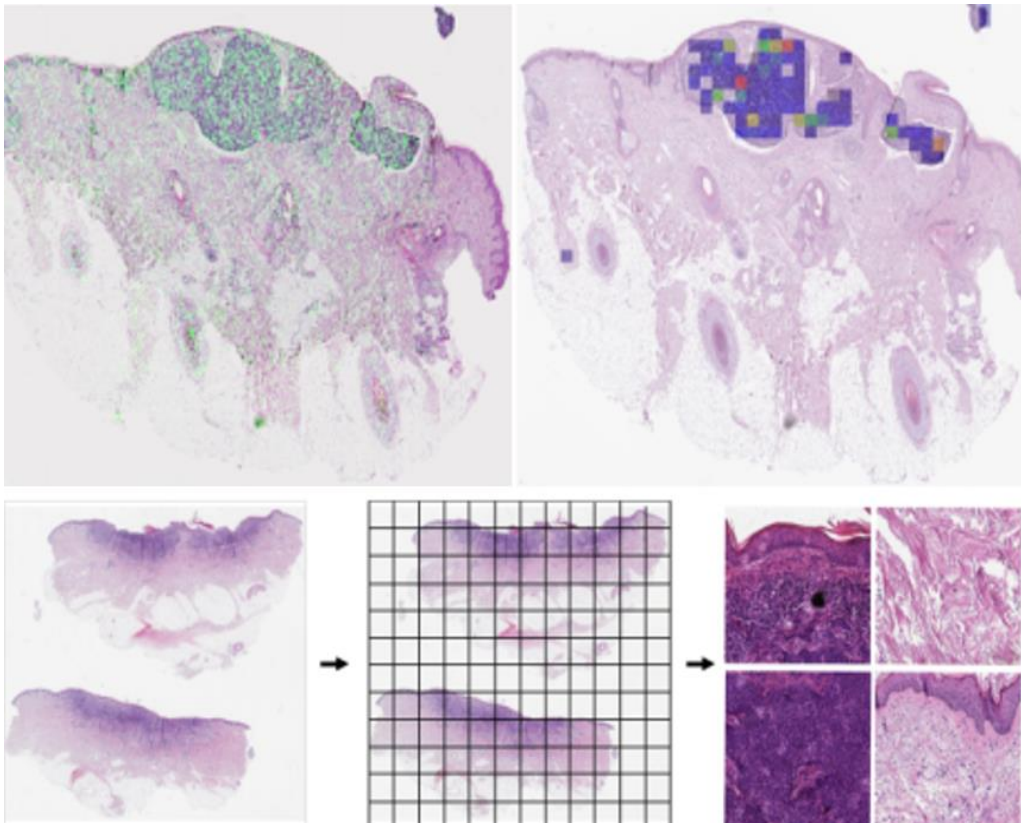


Figure 2: Methods of interpretive visualization. Green coloring indicates regions of the input image that contribute to the prediction using the Integrated Gradients method. Right: Regions of the input image that contribute to the prediction based on the attention mechanism's weights. Similar locations are identified as suggestive of the cancer class by both Integrated Gradients (left) and the attention mechanism (right).

## WSI method

All WSI were divided into 224224-pixel patches that covered the entire image, with padding applied where needed. The average color intensity  $c_p$  of all pixels for each patch  $p$  was determined to filter out empty background patches. The maximum  $c_{\max} = \max_p c_p$  was then determined for each WSI, and all patches with  $c_p$  more than 95% of this maximum were eliminated and considered empty. No stain-normalization was used, so each patch was normalized to zero mean and unit variance. The CNN was used to process each non-empty patch, and the activations of the last layer were saved as a representation of the patch. All patch activations of a single weakly supervised learning were sent to the network as a mini-batch, and these activations were subsequently used as input for the final classification network. We determined the mean or maximum of the generated network predictions using the mean and max pooling multi-instance learning techniques.

According to Ilse et al. (2018), the attention classifier was constructed so that the representation of each instance  $h_k$  corresponds to the network activations of one image patch. The weights are the result of a neural network that directly predicts these attention weights, thus we utilize a weighted average of instances.  $H = \{h_1, \dots, h_K\}$ , represents a set of  $z = \sum_{k=1}^K a_k h_k$  instances in a single entire slide image, and the MIL pooling function  $a_k$  is

$$a_k = \frac{\exp\{w^T \tanh(Vh_k^T)\}}{\sum_{j=1}^K \exp\{w^T \tanh(Vh_j^T)\}}$$

Where  $w \in \mathbb{R}^{L \times I}$  and  $V \in \mathbb{R}^{L \times I}$  are trained parameters. On a whole, a characterisation layer with sigmoid activation  $\sigma$  is functional to make available the estimate  $\hat{y} = \sigma(v^T z)$ . For training, we utilized regular SGD once more. The network was re-trained 100 times, with the mean and standard deviation of accuracy, F1 score, and AUC calculated each time.

## METHODS

Hematoxylin and eosin (H/E, n=838 slides) were used to stain images of normal skin and skin sections containing BCCs on histology slides. 647 of them are BCC cases, whereas 191 have normal skin. This batch of 838 photos was divided into 129 (15%) test images and 709 (85%) training images at random. The validation set was made up of 20% of the training set. The weakly supervised learning has a median size of 56,89626,198 pixels, with a height range of 6,884 to 47,939 pixels and a width range of 7,360 to 99,568 pixels. According to ethics decisions 1119/2018 (Ethics Committee Upper Austria) and 2085/2018 (Ethics Committee Upper Austria), the photos were retrospectively gathered at the Medical University of Vienna and the Kepler University Hospital (Ethics Committee Medical University of Vienna).

Using the PyTorch framework, we trained the baseline technique as well as three other ways based on mean pooling, max pooling, and the above-mentioned attention mechanism. The F1-scores were evaluated on various combinations in order to assess the impact of various instance-level and bag-level multi-instance learning pooling strategies. Because max pooling (Wang et al., 2018) and attention-based pooling (Ilse et al., 2018) have been attempted previously, this paper will test and report sum pooling, max pooling, attention-based pooling, and gated attention-based pooling.

## RESULTS AND DISCUSSIONS

Table 1 shows the outcomes of the attention-based approach versus the baseline method and mean/max pooling multi-instance learning on our data set. In fact, our method's F1-score surpassed all others. Despite our AUC and ROC scores, we outperformed the models by a considerable amount. The accuracy was slightly lower than the four traditional machine learning algorithms. In terms of AUC, the attention-based pooling technique outperformed multi-instance learning with max pooling, multi-instance learning with mean pooling, and an end-to-end CNN trained on down-sampled whole-slide images. Table 1 also includes the accuracy and F1 score of all predicting methods, resulting in the same ranking of the compared approaches. The multi-instance learning -attention approach has a sensitivity and specificity of 0.97.01 and 0.91.03, respectively. In regard to characterization, by evaluating the matching attention weights, this attention-based design makes it simple to discover the patches that are critical for network categorization (see Figure 2).

Table 1. Metrics of performance for the methods that were compared. The mean and mean±standard deviation across 100 re-runs of the training procedure are displayed. The Wilcoxon signed-rank test of AUC between CNN + MIL attention and the other approaches corresponds to the p-values.

Data type	Technique	Precision	F1 Score	AUC	P-value
patches	CNN + MIL attention	0.95 (0.94 – 0.95)	0.96 (0.96 -0.97)	0.99 (0.99 – 0.99)	
patches	CNN + MIL max pooling	0.91 (0.90 – 0.93)	0.95 (0.93 – 0.95)	0.97 (0.98 – 0.98)	<0.001
patches	CNN + MIL mean pooling	0.87 (0.86 – 0.89)	0.92 (0.91 -0.93)	0.92 (0.92 – 0.93)	<0.001
re-scaled WSI	End-to-end CNN	0.75 (0.69 – 0.81)	0.82 (0.76 – 0.87)	0.89 (0.88 – 0.91)	<0.001

Since they have been collected and maintained for decades, histopathology slides constitute a massive source of knowledge. Due to (1) the ultra-high resolution and (2) the weak labels, their computational processing provides a significant difficulty for machine learning algorithms. We've shown how to employ attention-based pooling and CNNs to detect basal cell carcinomas on histopathology slides, as well as how to evaluate and visualize those models.

## CONCLUSION

This lengthy summary describes ongoing and preliminary research. Other forms of attention mechanisms are investigated, as well as studies on new data sets. Because of its simplicity, we now extract features using the VGG11 architecture. We intend to compare this to different CNN architectures and train the entire pipeline from beginning to end. We also want to determine the best size-to-performance ratio by scaling down the weakly supervised learning step-by-step. Another option we intend to take is to compare the saliency maps collected by pathologists using eye tracking during diagnosis to the critical regions highlighted by attention weights.

## REFERENCES

- Abedin, M. M. M., Ahmed, A. A. A., and Neogy, T. K. (2012). Mechanism of Accountability and Auditing: Public Sector Scenarios of Bangladesh. *Journal of Business Studies*, 4, 131-148.
- Ahmed, A. A. A. & Dey, M. M. (2009a). Timeliness attributes and the extent of accounting disclosure: a study of banking companies in Bangladesh. *Osmania Journal of International Business Studies*, 4(1).

- Ahmed, A. A. A. (2009). The Effect of Timeliness Regulation of Corporate Financial Reporting: Evidence from Banking Sector of Bangladesh. *Accounting and Management Information Systems*, 8(2), 216 - 235. [http://online-cig.ase.ro/jcig/art/8\\_2\\_4.pdf](http://online-cig.ase.ro/jcig/art/8_2_4.pdf)
- Ahmed, A. A. A. and Neogy, T. K. (2009). Merger & Acquisitions (M&A) Goodwill Accounting: Principles and Practice. *The Bangladesh Accountant*, 65, 75-91.
- Ahmed, A. A. A., & Dey, M. M. (2009b). Corporate Attribute and the Extent of Disclosure: A Study of Banking Companies in Bangladesh. Proceedings of the 5th International Management Accounting Conference (IMAC), OCT 19-21, 2009, UKM, Kuala Lumpur, MALAYSIA, Pages: 531-553. <https://publons.com/publon/11427801/>
- Ahmed, A. A. A., Khan, W., & Hossain, M. S. (2011). Reporting Practice of Accounting Disclosure on Changes in Listed Companies of Bangladesh. *ASA University Review*, 5(1), 83-96. <https://www.researchgate.net/publication/336664901>
- Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S., and Navab, N. (2016). Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE transactions on medical imaging*, 35(5), 1313–1321.
- Amin, R., & Manavalan, M. (2017). Modeling Long Short-Term Memory in Quantum Optical Experiments. *International Journal of Reciprocal Symmetry and Physical Sciences*, 4, 6–13. Retrieved from <https://upright.pub/index.php/ijrsps/article/view/48>
- Andrews, S., Tsochantaridis, I. and Hofmann, T. (2003). Support vector machines for multiple-instance learning. In *Advances in neural information processing systems*, 577-584.
- Azad, M. R., Khan, W., & Ahmed, A. A. A. (2011). HR Practices in Banking Sector on Perceived Employee Performance: A Case of Bangladesh. *Eastern University Journal*, 3(3), 30–39. <https://doi.org/10.5281/zenodo.4043334>
- Begum, R., Ahmed, A. A. A., & Neogy, T. K. (2012). Management Decisions and Univariate Analysis: Effects on Corporate Governance in Bangladesh. *Journal of Business Studies*, 3, 87-115.
- Bynagari, N. B. (2018). On the ChEMBL Platform, a Large-scale Evaluation of Machine Learning Algorithms for Drug Target Prediction. *Asian Journal of Applied Science and Engineering*, 7, 53–64. Retrieved from <https://upright.pub/index.php/ajase/article/view/31>
- Bynagari, N. B. (2019). GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. *Asian Journal of Applied Science and Engineering*, 8, 25–34. Retrieved from <https://upright.pub/index.php/ajase/article/view/32>
- Bynagari, N. B., & Amin, R. (2019). Information Acquisition Driven by Reinforcement in Non-Deterministic Environments. *American Journal of Trade and Policy*, 6(3), 107-112. <https://doi.org/10.18034/ajtp.v6i3.569>
- Bynagari, N. B., & Fadziso, T. (2018). Theoretical Approaches of Machine Learning to Schizophrenia. *Engineering International*, 6(2), 155-168. <https://doi.org/10.18034/ei.v6i2.568>
- Chinem, V. P. and Miot, H. A. (2011). Epidemiology of basal cell carcinoma. *Anais Brasileiros de Dermatologia*, 86(2), 292–305.
- Dietterich, T. G., Lathrop, R. H. and Lozano-Pérez, T. (1997). Solving the multiple instance problem with axis-parallel rectangles. *Artificial intelligence*, 89(1-2), 31-71.
- Donepudi, P. K. (2014a). Technology Growth in Shipping Industry: An Overview. *American Journal of Trade and Policy*, 1(3), 137-142. <https://doi.org/10.18034/ajtp.v1i3.503>
- Donepudi, P. K. (2014b). Voice Search Technology: An Overview. *Engineering International*, 2(2), 91-102. <https://doi.org/10.18034/ei.v2i2.502>
- Donepudi, P. K. (2015). Crossing Point of Artificial Intelligence in Cybersecurity. *American Journal of Trade and Policy*, 2(3), 121-128. <https://doi.org/10.18034/ajtp.v2i3.493>



- Donepudi, P. K. (2016). Influence of Cloud Computing in Business: Are They Robust?. *Asian Journal of Applied Science and Engineering*, 5(3), 193-196. Retrieved from <https://journals.abc.us.org/index.php/ajase/article/view/1181>
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118.
- Fadziso, T., & Manavalan, M. (2017). Identical by Descent (IBD): Investigation of the Genetic Ties between Africans, Denisovans, and Neandertals. *Asian Journal of Humanity, Art and Literature*, 4(2), 157-170. <https://doi.org/10.18034/ajhal.v4i2.582>
- Foulds, J. and Frank, E., (2010). A review of multi-instance learning assumptions. *The Knowledge Engineering Review*, 25(1), 1-25.
- Griffin, J. and Treanor, D. (2017). Digital pathology in clinical use: Where are we now and what is holding us back? *Histopathology*, 70(1), 134–145.
- Ilse, M., Tomczak, J. M., and Welling, M. (2018). Attention-based deep multiple instance learning. arXiv preprint arXiv:1802.04712
- Janowczyk, A. and Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of pathology informatics*, 7.
- Komura, D. and Ishikawa, S. (2018). Machine learning methods for histopathological image analysis. *Computational and structural biotechnology journal*, 16, 34–42.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105.
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- Lin, Z., Feng, M., Santos, C. N. D., Yu, M., Xiang, B., Zhou, B. and Bengio, Y. (2017). A structured self-attentive sentence embedding. arXiv preprint arXiv:1703.03130.
- Manavalan, M. (2014). Fast Model-based Protein Homology Discovery without Alignment. *Asia Pacific Journal of Energy and Environment*, 1(2), 169-184. <https://doi.org/10.18034/apjee.v1i2.580>
- Manavalan, M. (2018). Do Internals of Neural Networks Make Sense in the Context of Hydrology?. *Asian Journal of Applied Science and Engineering*, 7, 75–84. Retrieved from <https://upright.pub/index.php/ajase/article/view/41>
- Manavalan, M. (2019). P-SVM Gene Selection for Automated Microarray Categorization. *International Journal of Reciprocal Symmetry and Physical Sciences*, 6, 1–7. Retrieved from <https://upright.pub/index.php/ijrsps/article/view/43>
- Manavalan, M. (2019). Using Fuzzy Equivalence Relations to Model Position Specificity in Sequence Kernels. *Asian Journal of Applied Science and Engineering*, 8, 51–64. Retrieved from <https://upright.pub/index.php/ajase/article/view/42>
- Momeni, A., Thibault, M., and Gevaert, O. (2018). Deep recurrent attention models for histopathological image analysis. *BioRxiv*, page 438341.
- Neogy, T. K. and Ahmed, A. A. A. (2015). The Extent of Disclosure of Different Components of Disclosure Index: A Study on Commercial Banks in Bangladesh. *Global Disclosure of Economics and Business*, 4(2), 100-110. <https://doi.org/10.18034/gdeb.v4i2.139>
- Qi, C. R., Su, H., Mo, K. and Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. *Proc. Computer Vision and Pattern Recognition (CVPR)*, IEEE, 1(2), p.4.
- Ramon, J. and De Raedt, L. (2000). Multi instance neural networks. In *Proceedings of the ICML-2000 workshop on attribute-value and relational learning*, 53-60.

- Siddique, M. N. & Ahmed, A. A. A. (2015). Congruence of Competitive Advantage and Transfer Pricing: A Study on Selected MNCs Operating in Bangladesh. *Asian Accounting & Auditing Advancement*, 5(2), 119-126. <https://www.researchgate.net/publication/354712086>
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
- Sun, X., Tan, J., Tang, L., Guo, J. J. and Li, X. (2018). Real world evidence: experience and lessons from China. *Bmj*, 360, p.j5262.
- Sun, X., Tan, J., Tang, L., Guo, J. J. and Li, X. (2018). Real world evidence: experience and lessons from China. *Bmj*, 360, p.j5262.
- Sundararajan, M., Taly, A., and Yan, Q. (2017). Axiomatic Attribution for Deep Networks. In ICML'17 Proceedings of the 34th International Conference on Machine Learning - Volume 70, 3319-3328.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems, 5998-6008.
- Wang, X., Yan, Y., Tang, P., Bai, X. and Liu, W. (2018). Revisiting multiple instance neural networks. *Pattern Recognition*, 74, 15-24.
- Wang, Z., Poon, J., Sun, S. and Poon, S., 2018. CNN based Multi-Instance Multi-Task Learning for Syndrome Differentiation of Diabetic Patients. arXiv preprint arXiv:1812.07764.
- Wei, X. S., Wu, J. and Zhou, Z. H. (2014). Scalable multiinstance learning. In Data Mining (ICDM), 2014 IEEE International Conference, 1037-1042.
- Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R. and Bengio, Y. (2015). Show, attend and tell: Neural image caption generation with visual attention. In International conference on machine learning, 2048-2057.
- Yan, Y., Wang, X., Guo, X., Fang, J., Liu, W. and Huang, J. (2018). Deep Multi-instance Learning with Dynamic Pooling. In Asian Conference on Machine Learning, 662-677.
- Zhang, M. L. and Zhou, Z. H. (2004). Improve multi-instance neural networks through feature selection. *Neural Processing Letters*, 19(1), 1-10.
- Zhang, Q. and Goldman, S. A. (2002). EM-DD: An improved multipleinstance learning technique. In Advances in neural information processing systems, 1073-1080.
- Zhao, Y., He, L., Xie, Q., Li, G., Liu, B., Wang, J., Zhang, X., Zhang, X., Luo, L., Li, K. and Jing, X. (2015). A novel classification method for syndrome differentiation of patients with AIDS. Evidence-Based Complementary and Alternative Medicine.
- Zhou, Z. H. (2017). A brief introduction to weakly supervised learning. *National Science Review*, 5(1), 44-53.
- Zhou, Z. H., Sun, Y. Y. and Li, Y. F. (2009). Multi-instance learning by treating instances as non-iid samples. In Proceedings of the 26<sup>th</sup> annual international conference on machine learning (pp. 1249-1256).