IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 14, 2020, accepted October 29, 2020, date of publication November 5, 2020, date of current version November 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3036090

SpineNet-6mA: A Novel Deep Learning Tool for Predicting DNA N6-Methyladenine Sites in Genomes

ZEESHAN ABBAS[1,2], **HILAL TAYARA**[3], **AND KIL TO CHONG**[1,4], **(Member, IEEE)**

[1]Department of Electronics and Information Engineering, Jeonbuk National University, Jeonju 54896, South Korea
[2]Institute of Avionics and Aeronautics (IAA), Air University, Islamabad 44000, Pakistan
[3]School of International Engineering and Science, Jeonbuk National University, Jeonju 54896, South Korea
[4]Information Research Center, Jeonbuk National University, Jeonju 54896, South Korea

Corresponding authors: Kil To Chong (kitchong@jbnu.ac.kr) and Hilal Tayara (hilaltayara@jbnu.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant Funded by the Korea Government Ministry of Science and ICT (MSIT) under Grant 2020R1A2C2005612, and in part by the Brain Research Program of the National Research Foundation (NRF) Funded by the Korean Government (MSIT) under Grant NRF-2017M3C7A1044816.

ABSTRACT DNA N6-methyladenine (6mA) has subsequently been identified as an important epigenetic modification which plays an important role in various cellular processes. The precise discrimination of N6-methyladenine (6mA) in genomes is required to recognize its biological functions. Although, we have several experimental techniques for the identification of 6mA-sites, *in silico* prediction has evolved as an alternative approach due to high-cost and labor-intense in experimental techniques. Taking into account, the implementation of an efficient and accurate model for identification of N6-methyladenine is of high priority. Several machine learning and deep learning models have already been developed to classify genome-wide 6mA sites. However, their success in predicting 6mA sites still has room for improvement. Based on this, we proposed a novel deep learning based model for the prediction of DNA N6-methyladenine sites in rice genomes. We built our model based on a special architecture called SpinalNet using DNA 6mA sites in rice genome and obtained an accuracies of 94.31% and 94.77% with an MCCs of 0.88 and 0.89 on two different datasets. The model generalizes well to other genomes as well, validated through cross-species testing. The results validate that the proposed model produces better scores than existing models regarding all evaluation parameters. A user-friendly webserver is made available at http://nsclbio.jbnu.ac.kr/tools/SpineNet6mA/.

INDEX TERMS Deep learning, DNA sequence, epigenetics, neural networks, spinalnet.

I. INTRODUCTION

DNA N6-methyladenine (6mA) is an important epigenetic modification of diverse species genomes, found in bacteria, eukaryotes, and archaea [1], [2]. It refers to the methylation at the 6th position of an adenine ring and is a highly researched subject in epigenetics [3]–[6]. DNA 6mA modification is usually the one of the most widespread DNA modification in prokaryotic genomes, once thought to be non-existent in eukaryotes (including human-being), since it was not found in earlier studies [7]. This process plays an important role in regulating various biological processes, including the system of restriction-modification, cell defense, DNA repair and replication, and gene expression [8]. Studying the distribution of DNA 6mA may provide a bottomless understanding of the process of epigenetic modification. Modern studies have shown that the abnormal state of modification of DNA 6mA is linked to human cancer and other diseases [9]. Numerous 6mA modifications have been observed in various multicellular eukaryotes with the advent of high-throughput sequencing technology, but still inadequate. The development of experimental techniques contributes to the analysis of 6mA modifications. *In silico* prediction of DNA 6mA sites in a genome have evolved as an alternative approach due to the constraint of labor-intensive and costly experiments.

Recently, the initiation of experimental approaches using the machine and deep learning methods have overwhelmed numerous complications in recognizing 6mA modifications. The 6mA modification has always been a hot topic in research, and a lot of researchers are using the machine and

The associate editor coordinating the review of this manuscript and approving it for publication was Tossapon Boongoen.

deep learning algorithms to recognise 6mA sites in the rice genome [10]–[12].

Feature extraction always plays a significant role in machine learing during the creation of any predictor [13]–[16]. Researchers use different feature extraction algorithms like binary encoding, nucleotide chemical properties, KMER, and Markov features, etc. in their models. A machine learning based tool using support vector machine (SVM) named iDNA6mA-PseKNC was proposed by Feng *et al.,* to anticipate 6mA sites in Mus musculus genome unveiling that this method is tried and tested to recognize genome-wide 6mA sites in numerous species [17]. Lately, Chen *et al.,* provided a standard dataset, 6mA-rice-Chen, for 6mA prediction containing equal number of 6mA and non-6mA sites, 880 each, in the rice genome [18]. To ascertain 6mA sites in the rice genome, they built an SVM based tool (i6mA-Pred) using many handcrafted DNA sequence features. On this dataset, 83% accuracy was claimed using the i6mA-Pred tool. Using the same benchmark dataset Pian *et al.,* made a prediction of 6mA sites using their model MM-6mAPred based on the Markov model and outperformed the i6mA-Pred tool [19]. Following the same, Tahir *et al.,* proposed another model named iDNA6mA for the same purpose of identification of 6mA in the rice genome. They again trained and tested their model on the same dataset used by Chen *et al.,* and outperformed i6mA-Pred while predicting 6mA sites [10]. Basith *et al.,* proposed a model using ensemble approach for the prediction of rice genome. They used several feature encoding methodologies and machine learning classifiers and named the model, SDM6A. They claimed that SDM6A outperformed the previous models including i6mA-Pred and iDNA6mA on the same benchmark dataset [20]. Lv *et al.,* came up with another benchmark dataset with a huge number of sequences for rice genome named 6mA-rice-Lv and proposed their tool, iDNA6mA-rice. The dataset 6mA-rice-Lv consists of 154,000 positive sequences and 154,000 negative ones, where the positive sequences contain the 6mA sites. They trained and tested iDNA6mA-rice on this dataset using five-fold cross-validation and achieved good results [21]. Yu and Dai proposed a model called SNNRice6mA based on convolutional neural network (CNN) for improving the prediction accuracy using 6mA-rice-Lv dataset [22]. Using the same five-fold cross-validation SNNRice6mA achieved an accuracy of 92.02% on 6mA-rice-Lv dataset which is the better among all of the previous methodologies.

Deep Neural Networks (DNNs) have brought state-of-the-art success in various fields of science and engineering [11], [12], [23], [24]. Earlier studies demonstrated that deep learning is an impressive technique for analyzing and classifying sequences in bioinformatics [25]–[32]. Current DNNs have vanishing gradient problem, as the number of layers increases, the number of connections increases persistently. Narrowing down to few neurons is unfeasible in current DNN architectures if we have a large number of inputs. In [33] Kabir *et al.,* proposed a new DNN architecture called SpinalNet to overcome the above-mentioned issues. The SpinalNet architecture has been explained in Section II.

In this article, we have proposed a deep architecture using SpinalNet for the first time for sequence data to get high accuracy while predicting the 6mA sites in rice genome. Unlike traditional machine learning methods, it learns high-level abstract features by using the CNN architecture. The architecture has been trained and validated on 6mA-rice-Lv and 6mA-rice-Chen benchmark datasets and achieved an accuracy of 94.31% and 94.77% with an MCC of 0.8868 and 0.8966 respectively and outperformed all the previous methods. The main contributions in this study are providing a novel deep learning architecture for bioinformatics applications, in addition, improving the state-of-the-art performance of the 6mA computational models.

This article is organized as follows: Section II presents the SpinalNet network, Section III introduces the datasets used in this study, Section IV presents the proposed methodology, Section V introduces the evaluation metrics used to study the performance of the proposed model, Section VI presents the achieved results, Section VII shows the performance of the proposed design on different species, and Section VIII concludes the paper.

II. SPINALNET

By imitating the functionality of the visual cortex of cats, convolutional neural networks (CNN) have been developed by the researchers which dramatically improved the NNs accuracy [34]. By observing the recent success of CNNs and the wonderful architecture of human spinal cord Kabir *et al.,* developed a NN with gradual inputs, named SpinalNet as shown in Figure 1 [33].

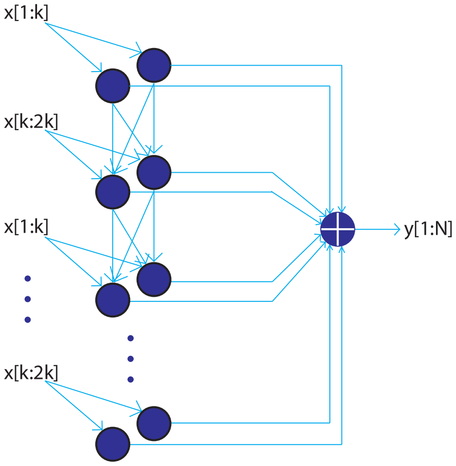
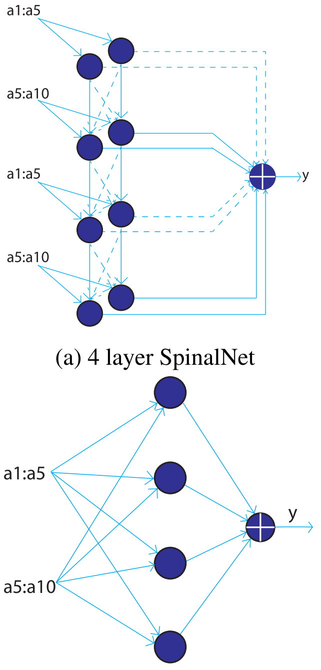


FIGURE 1. Structure of SpinalNet.

The configuration of the network is composed of input row, an intermediate row of multiple hidden layers, and the output row. To minimize the number of multiplications, we keep the number of inputs per layer and the number of neurons per hidden layers as small as possible but this may cause the network to underfit. To overcome this issue, each layer in the intermediate row receives input from the previous layer. Since the input is recurring, if any significant input feature does not affect the output in one of the hidden layers, they can affect in another hidden layer. The input is split into two rows in Figure 1 and both rows are allocated to different hidden layers repetitively.

A generalized version of SpinalNet having 4 hidden layers containing 2 neurons per layer is shown in Figure 2a. The first layer of SpinalNet just takes the weighted sum of inputs a1 to a5. The outputs of the hidden neurons from the first hidden layer go only to the related neurons of the second hidden layer. By assigning zero weight, interconnections and connections from the very first hidden layer to the output are disconnected. The following hidden layer is given the weighted sum of a6 to a10 along with the weighted sum of the previous layer. So this layer's neurons apply an activation function to the weighted sum of previous layers inputs a1 to a10. A similar approach will be applied to the next two layers. Therefore, the SpinalNet having 4 hidden layers containing two neurons per layer is equivalent to a traditional neural network of one hidden layer with 4 neurons as shown in Figure 2b.



(a) 4 layer SpinalNet

(b) 1 layer traditional NN

FIGURE 2. 4 layer SpinalNet equivalent to 1 layer traditional neural network.

III. DATASETS

In this study, we have used the 6mA-rice-Lv dataset, which was previously being used by [21] and [22] in their deep learning models. The sequences in this dataset were experimentally identified in the study carried out by Zhou *et al.,* [35]. These identified sequences were deposited in the NCBI Gene Expression Omnibus https://www.ncbi.nlm.nih.gov/geo/ under the accession number GSE103145. It contains a total of 265,290 6mA sites. Thus, to avoid redundancy and remove the homologous bias, CD-HIT program was used to remove the similar sequences with the similarity above 80%. As a result, 154,000 6mA sites-contained sequences were selected as positive dataset. Negative dataset was prepared from the NCBI database https://www.ncbi.nlm.nih.gov/genome/10 by randomly selecting 154,000 sequences that have not-methylated-adenine in the center and rich of GAGG, AGG, and AG motifs. These motifs are reported to be frequent in 6mA sequences [35].

We considered another benchmark dataset, 6mA-rice-Chen, comprises of 880 each of positive and negative sample sequences. This dataset has been used by many researchers to evaluate their models [10], [18]–[21]. Along with these, we have used two another datasets, 6mA-chinensis and 6mA-vesca, for cross-species testing. In both positive and negative samples, the DNA sequences are 41bp long. All the positive samples had a 6mA site in the middle, while the center for each negative sequence has no 6mA change site. Table 1 presents an in depth view of the datasets:

TABLE 1. Details of benchmark datasets used.

Dataset	Positive	Negative	Total	Species
6mA-rice-Lv	154,000	154,000	308,000	rice
6mA-rice-Chen	880	880	1760	rice
6mA-chinensis	813	813	1626	rosaceae
6mA-vesca	1966	1966	3932	rosaceae

IV. PROPOSED METHODOLOGY

The samples in the used dataset are sequences of DNA, represented as a string like *AGTACT* ... *CAT*. Since neural networks accept only numerical data we need to represent the strings first into a form that is acceptable to the network. Therefore, we have used binary encoding which is an effective scheme to convert the nucleotides into such format:

$$A : 1, 0, 0, 0$$
$$T : 0, 1, 0, 0$$
$$C : 0, 0, 1, 0$$
$$G : 0, 0, 0, 1$$

Here, the nucleotide representation is not novel, and depictions of A, T, C, and G are exchangeable. The binary encoding algorithm can be interpreted as a single nucleotide representation also called a one-hot-encoding algorithm. It is then possible to transform a random DNA sequence containing m nucleotides into a matrix of features having shape $4 \times m$ [21], [36].

We built a novel deep learning model SpineNet-6mA using SpinalNet. SpineNet-6mA consists of five convolutional layers followed by pooling and dropout layers. The convolution greatly surges the number of parameters. Pooling may be

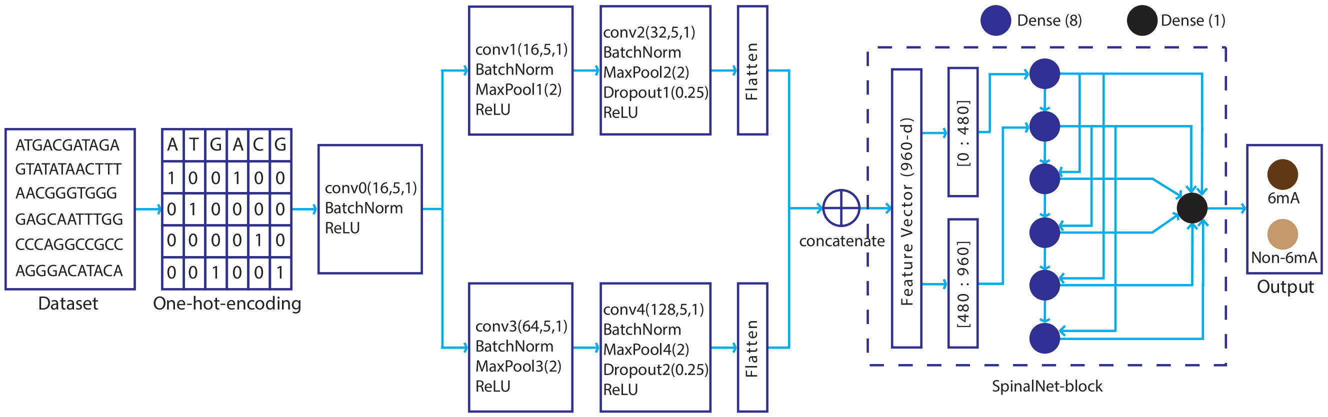


FIGURE 3. SpineNet-6mA architecture for the identification of DNA 6mA modification.

applied to minimize the number of parameters by reducing the size of input for the next layer but it also causes loss of information [37]. The input to the model is the binary encoded sequences. The block diagram of the proposed model is as shown in Figure 3.

Convolution layer [38] named conv0 having 16 filters with a filter size of 5 is applied on the input matrix followed by batch normalization [39] named normLayer0 and activation layer named act0. ReLU [40] has been used as a non-linear activation function throughout the network, except sigmoid on the output layer. Mathematically, ReLU, and sigmoid can be expressed as:

$$ReLU(x) = max(0, x)$$
$$Sigmoid(x) = \frac{1}{1 + e^{-x}}$$

The model from here is divided into two sub-branches, the first branch consists of conv1 comprises of 16 filters again of size 5 each followed by normLayer1 which is followed by pool1 of size 2, and then act1. The output of act1 is given as input to conv2 having 32 filters of size 5 each followed by normLayer2 and pool2 of the same size above. It is then followed by a dropout layer named dropoutLayer1 with a dropout value of 0.25 which is followed by the activation layer act2. The dropout process discards some intermediate features and improves the reliability of the model by preventing the overfitting problem. The second branch is just similar to the first one except for the difference in the number of filters in convolution layers. The first convolution layer in the second branch named conv3 consist of 64 filters and the second convolution layer named conv4 consist of 128 filters. After the activation layers in both branches, the outputs are flattened respectively to integrate the intermediate features and then concatenated to get a single flattened output. The dimensionality reduction is a common technique to reduce the number of inputs to a neural network without significant degradation of output [41]. The network's input combination may include a large number of interrelated and unrelated data. Therefor, the concatenated flattened output is then fed to the SpinalNet-block having 6 dense layers of 8 neurons in each dense layer. The input to the SpinalNet-block is the 960-d vector which is the concatenation of the learned features from the two CNN branches of the model. This vector is then split into two vectors of 480-d for both to be used in the spinal net block as shown in Figure 3. Since SpinalNet has fewer neurons per hidden layer compared to the number of inputs, it takes input in every layer and may automatically repudiate insignificant data. The output of the dense layers is then fed to the sigmoid function for the classification purpose. The sigmoid function provides a float value from 0 to 1, which is known to be the likelihood of the 6mA change site in the input DNA sequence. If the likelihood is greater than 0.5, the model classifies the sequence as positive, and if it is less than 0.5 the model classifies the sequence as negative [42]. The positive sample indicates the center of the input DNA sequence as the 6mA site.

We used stochastic gradient descent (SGD) as an optimizer by setting the learning rate as 0.001 and a momentum of 0.95. Furthermore, we have used a learning rate scheduler during the training phase to reduce the learning rate if the loss value on the validation set is not reduced anymore. For the first ten epochs, the scheduler preserves the initial learning rate and then reduces it exponentially. The maximum epochs for training are set to 100 and the training batch size to 32. In addition, we used early stopping [43] with the patience of 30 epochs that specifies, it will stop the training process when the prediction accuracy on the validation set stops improving after 30 epochs.

We put our model into effect based on Keras 2.3.1. The ideal hyper-parameters are observed by making use of the well-known grid search algorithm. The used hyper-parameter values have been mentioned in Table 2.

V. PERFORMANCE EVALUATION METRICS

We used the traditional 5-fold and 10-fold cross-validation method to validate our approach to be compatible with the

TABLE 2. Hyper-parameters used in SpineNet-6mA.

Hyper-parameters	Values
Number of filters	16,16,32,64,128
Kernel size (convolution)	5
Kernel size (pooling)	2
Neurons per hidden layer	8
Dropout rate	0.25
Learning rate	0.001
Momentum rate	0.95
Batch size	32

previous studies using the benchmark datasets 6mA-rice-Lv and 6mA-rice-Chen. To evaluate the performance, we utilized the same five metrics including sensitivity, specificity, accuracy, Matthews correlation coefficient (MCC), and area under the curve (AUC) to remain consistent with the previous methodologies.

The metric sensitivity is also known as True Positive Rate (TPR), can be expressed as:

$$Sensitivity = TPR = \frac{TP}{TP + FN}$$

where, TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative, respectively.

Specificity is also known as True Negative Rate (TNR) and expressed as:

$$Specificity = TNR = \frac{TN}{TN + FP}$$

The correct predictions to the test data is called accuracy. It can be defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

The Matthews correlation coefficient (MCC) reflects the model's output as a binary classifier [44]. It can be defined as:

$$MCC = \frac{TP \times \text{TN-FP} \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

The last figure of merit we used is the AUC defined as the area under the receiver operating curve (ROC). The value of AUC ranges between 0 and 1, where 1 indicates the perfection of the model.

VI. RESULTS

Proving that our methodology SpineNet-6mA is superior to other approaches using the benchmark dataset 6mA-rice-Lv, we contrasted it with the two existing state-of-the-art tools including iDNA6mA-rice [21] and SNNRice6ma-large [22] to recognize DNA 6mA sites in rice genome. Results showed better performance of our method, SpineNet-6mA, than those of the above-mentioned tools.

The two state-of-the-art techniques to which we are comparing our results on the benchmark dataset, 6mA-rice-Lv, used 5-fold cross-validation to obtain their results. To achieve a better comparative analysis, we used the same validation strategy to validate SpineNet-6mA by keeping the same number of folds. We make use of the corresponding five metrics including sensitivity, specificity, accuracy, MCC and AUC, to remain aligned with the assessment metrics used in these studies. From the original study, the outputs of iDNA6mA-rice [21] and SNNRice6mA-large [22] has been quoted directly in Table 3. We found that SpineNet-6mA outperformed both iDNA6mA-rice and SNNRice6ma-large in all five evaluation metrics. A comparison of the proposed model with the above-mentioned existing techniques using 6mA-rice-Lv dataset is shown in Table 3. From Table 3 we can see that SpineNet-6mA outperforms the model iDNA6mA-rice by 2.71% of sensitivity, 2.42% of specificity, 2.61% of accuracy, 4% of MCC, and 2% of AUC. Similarly, the SpineNet-6mA outperforms the model SNNRice6mA-large by 1.38% of sensitivity, 3.17% of specificity, 2.27%of accuracy, 4% of MCC, and 1% of AUC. The graphical illustration of the experimental results is shown in Figure 4.

TABLE 3. 5-fold cross-validation performance comparison between iDNA6mA-rice, SNNRice6mA-large, and SpineNet-6mA on 6mA-rice-Lv dataset.

Methods	Sensitivity (%)	Specificity (%)	Accuracy (%)	MCC	AUC
iDNA6mA-rice	93.00	90.50	91.70	0.84	0.96
SNNRice6mA-large	94.33	89.75	92.04	0.84	0.97
SpineNet-6mA	95.71	92.92	94.31	0.88	0.98

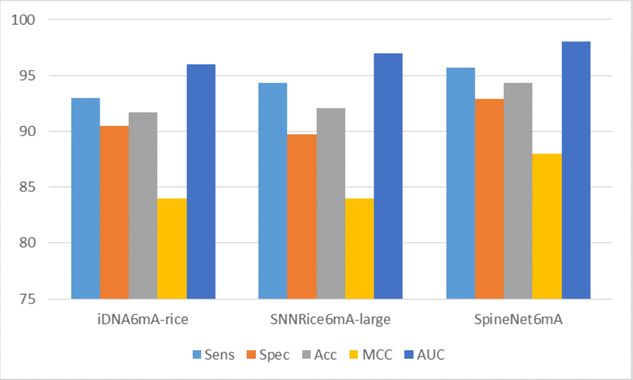


FIGURE 4. 5-fold cross-validation performance comparison between iDNA6mA-rice, SNNRice6mA-large, and SpineNet-6mA on 6mA-rice-Lv dataset, Graphical Representation.

The achieved results by SpineNet-6mA are stable with very low standard deviations. They are given as, sensitivity: 95.71 ± 0.007%, specificity: 92.92 ± 0.015%, accuracy: 94.31 ± 0.009%, MCC: 88.68 ± 0.018%, and AUC: 98.19 ± 0.005%.

The receiver operating characteristic curve of SpineNet-6mA is shown below in Figure 5, and comparison of the ROC curves between SpineNet-6mA, SNNRice6mA-large,

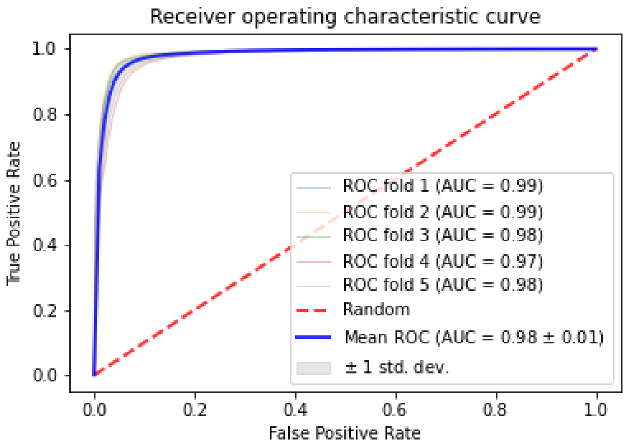


FIGURE 5. ROC curve of SpineNet-6mA on 6mA-rice-Lv dataset.

and iDNA6mA-rice on the testing set of 6mA-rice-Lv dataset is shown below in Figure 6.

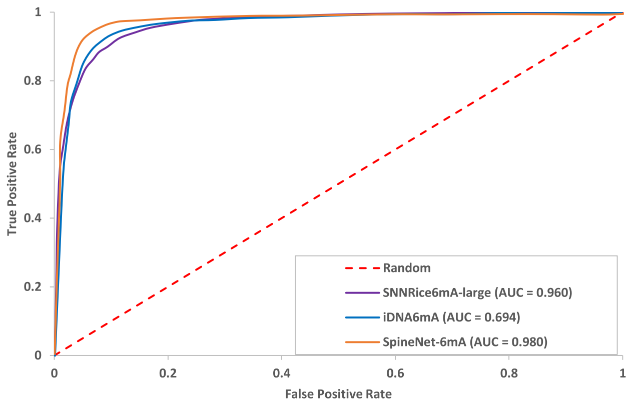


FIGURE 6. ROC curves comparison of SpineNet-6mA with SNNRice6mA-large and iDNA6mA on 6mA-rice-Lv dataset.

To show the robustness of our model, we used one small dataset of rice, 6mA-rice-Chen, and compared the results with the previous well-known models. The existing models for the identification of 6mA sites in rice genome includes, i6mA-Pred [18], iDNA6mA [10], SDM6A [20], MM-6mAPred [19], iDNA6mA-rice [21] and DNA6mA-MINT [26]. To be consistent and to have a fair comparison we trained and tested the model using 10-fold cross-validation on 6mA-rice-Chen dataset and evaluated the same five metrics, specificity, sensitivity, accuracy, MCC, and AUC.

From Table 4 we can clearly see that SpineNet-6mA outperforms all previous methodologies with respect to all evaluation metrics.

TABLE 4. 10-fold cross-validation performance comparison between SpineNet-6mA and previous tools on 6mA-rice-Chen dataset.

Methods	Sensitivity (%)	Specificity (%)	Accuracy (%)	MCC	AUC
iDNA6mA-rice	83.86	83.41	83.63	0.67	0.91
SDM6A	85.20	90.90	88.10	0.76	0.94
iDNA6mA	86.70	86.59	86.64	0.73	0.93
MM-6mAPred	89.32	90.11	89.72	0.79	-
i6mA-Pred	82.95	83.30	83.13	0.66	0.89
SNNRice6mA	92.16	94.32	93.24	0.87	0.97
DNA6mA-MINT	94.25	90.80	92.53	0.85	0.95
SpineNet-6mA	93.75	95.79	94.77	0.89	0.98

VII. CROSS-SPECIES EVALUATION

To check the model's validity which is trained on rice-dataset, we tested on other species Rosa chinensis [45] and Fragaria vesca [46] to predict DNA 6mA sites. We denoted these datasets as 6mA-chinensis and 6mA-vesca. 6mA-chinensis contains 813 positive and 813 negative sequences, while, 6mA-vesca contains 1966 positive and 1966 negative sequences. We also performed similar testing on the two other methods iDNA6mA-rice and SNNRice6ma-large using these two datasets and found that SpineNet-6mA has achieved higher performance when compared with the existing tools. For SNNRice-6mA-large they have provided the weights so using those weights we performed the testing while iDNA6ma-rice has provided their online server for testing purposes. Using the server we calculated TP, TN, FP, and FN and then calculated the accuracy and MCC using equations provided in section V. A comparison of the results between SpineNet-6mA, iDNA6mA-rice, and SNNRice6mA-large on Rosa chinensis [45] and Fragaria vesca [46] datasets can be seen as in Table 5. SpineNet-6mA outperformed the models

TABLE 5. Cross-Species Performance comparison between iDNA6mA-rice, SNNRice6mA-large, and SpineNet-6mA on 6mA-chinensis and 6mA-vesca datasets.

Methods	6mA-chinensis		6mA-vesca	
	Accuracy (%)	MCC	Accuracy (%)	MCC
iDNA6mA-rice	83.15	0.67	87.07	0.75
SNNRice6mA-large	81.13	0.62	87.84	0.75
SpineNet-6mA	85.20	0.70	90.30	0.80

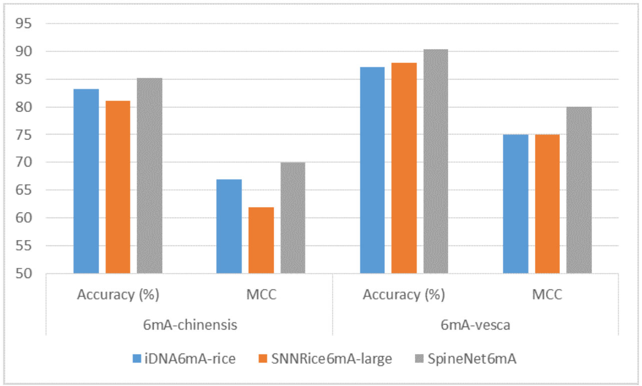


FIGURE 7. Cross-species performance comparison between iDNA6mA-rice, SNNRice6mA-large, and SpineNet-6mA on 6mA-chinensis and 6mA-vesca datasets, Graphical Representation.

iDNA6mA and SNNRice6mA in accuracy by 2.05% and 4.07% on 6mA-chinensis dataset while 3.23% and 2.46% on 6mA-vesca dataset respectively.

The graphical illustration of the cross-species performance results is shown in Figure 7.

VIII. CONCLUSION

Accuracy in the detection of N6-methyladenine (6mA) is of great importance. In this study, we developed a novel deep learning model SpineNet-6mA for the detection of 6mA sites with high accuracy. The proposed model is based on a special architecture called SpinalNet which attempts to imitate the human somatosensory system to effectively receive large data and achieve better efficiency. Using this model, we achieved an accuracy of 94.31% and 94.77% which is 2.27% and 1.53% better than the best existing state-of-the-art models on 6mA-rice-Lv and 6mA-rice-Chen datasets respectively. The proposed model also produced state-of-the-art results on testing another species not used in the training. We reckon that our proposed method can indeed be genuinely beneficial for the detection of 6mA-sites and hence be helpful in drug discovery and the bioinformatics field. For good, SpineNet-6mA has been made available at http://nsclbio.jbnu.ac.kr/tools/SpineNet6mA/ for free access.

REFERENCES

[1] Z. K. O'Brown and E. L. Greer, "N6-methyladenine: A conserved and dynamic dna mark," in *DNA Methyltransferases-Role and Function*. Cham, Switzerland: Springer, 2016, pp. 213–246.

[2] G. Zhang, H. Huang, D. Liu, Y. Cheng, X. Liu, W. Zhang, R. Yin, D. Zhang, P. Zhang, J. Liu, C. Li, B. Liu, Y. Luo, Y. Zhu, N. Zhang, S. He, C. He, H. Wang, and D. Chen, "N6-methyladenine DNA modification in drosophila," *Cell*, vol. 161, no. 4, pp. 893–906, May 2015.

[3] G.-Z. Luo, M. A. Blanco, E. L. Greer, C. He, and Y. Shi, "Dna N^6-methyladenine: A new epigenetic mark in eukaryotes?" *Nature Rev. Mol. Cell Biol.*, vol. 16, no. 12, pp. 705–710, 2015.

[4] B. Liu, F. Weng, D.-S. Huang, and K.-C. Chou, "IRO-3wPseKNC: Identify DNA replication origins by three-window-based PseKNC," *Bioinformatics*, vol. 34, no. 18, pp. 3086–3093, Sep. 2018.

[5] A. Wahab, S. D. Ali, H. Tayara, and K. T. Chong, "iIM-CNN: Intelligent identifier of 6mA sites on different species by using convolution neural network," *IEEE Access*, vol. 7, pp. 178577–178583, 2019.

[6] W. Alam, S. D. Ali, H. Tayara, and K. Chong, "A CNN-based RNA N6-methyladenosine site predictor for multiple species using heterogeneous features representation," *IEEE Access*, vol. 8, pp. 138203–138209, 2020.

[7] J. Casadesús and D. Low, "Epigenetic gene regulation in the bacterial world," *Microbiol. Mol. Biol. Rev.*, vol. 70, no. 3, pp. 830–856, Sep. 2006.

[8] W. He, C. Jia, Y. Duan, and Q. Zou, "70ProPred: A predictor for discovering sigma70 promoters based on combining multiple features," *BMC Syst. Biol.*, vol. 12, no. S4, p. 44, Apr. 2018.

[9] C.-L. Xiao, S. Zhu, M. He, D. Chen, Q. Zhang, Y. Chen, G. Yu, J. Liu, S. Q. Xie, F. Luo, and Z. Liang, "N^6-methyladenine DNA modification in the human genome," *Mol. cell*, vol. 71, no. 2, pp. 306–318, 2018.

[10] M. Tahir, H. Tayara, and K. T. Chong, "IDNA6 mA (5-step rule): Identification of DNA N6-methyladenine sites in the rice genome by intelligent computational model via Chou's 5-step rule," *Chemometric Intell. Lab. Syst.*, vol. 189, pp. 96–101, Jun. 2019.

[11] A. Byerly, T. Kalganova, and I. Dear, "A branching and merging convolutional network with homogeneous filter capsules," 2020, *arXiv:2001.09136*. [Online]. Available: http://arxiv.org/abs/2001.09136

[12] K. Kowsari, M. Heidarysafa, D. E. Brown, K. J. Meimandi, and L. E. Barnes, "RMDL: Random multimodel deep learning for classification," in *Proc. 2nd Int. Conf. Inf. Syst. Data Mining (ICISDM)*, 2018, pp. 19–28.

[13] Y. Wang, S. Yang, J. Zhao, W. Du, Y. Liang, C. Wang, F. Zhou, Y. Tian, and Q. Ma, "Using machine learning to measure relatedness between genes: A multi-features model," *Sci. Rep.*, vol. 9, no. 1, pp. 1–15, Dec. 2019.

[14] N. Stephenson, E. Shane, J. Chase, J. Rowland, D. Ries, N. Justice, J. Zhang, L. Chan, and R. Cao, "Survey of machine learning techniques in drug discovery," *Current Drug Metabolism*, vol. 20, no. 3, pp. 185–193, May 2019.

[15] J. Song, Y. Wang, F. Li, T. Akutsu, N. D. Rawlings, G. I. Webb, and K.-C. Chou, "IProt-sub: A comprehensive package for accurately mapping and predicting protease-specific substrates and cleavage sites," *Briefings Bioinf.*, vol. 20, no. 2, pp. 638–658, Mar. 2019.

[16] L. Wei, C. Zhou, H. Chen, J. Song, and R. Su, "ACPred-FL: A sequence-based predictor using effective feature representation to improve the prediction of anti-cancer peptides," *Bioinformatics*, vol. 34, no. 23, pp. 4007–4016, Jun. 2018.

[17] P. Feng, H. Yang, H. Ding, H. Lin, W. Chen, and K.-C. Chou, "IDNA6mA-PseKNC: Identifying DNA N6-methyladenosine sites by incorporating nucleotide physicochemical properties into PseKNC," *Genomics*, vol. 111, no. 1, pp. 96–102, Jan. 2019.

[18] W. Chen, H. Lv, F. Nie, and H. Lin, "I6mA-pred: Identifying DNA N^6-methyladenine sites in the rice genome," *Bioinformatics*, vol. 35, no. 16, pp. 2796–2800, Aug. 2019.

[19] C. Pian, G. Zhang, F. Li, and X. Fan, "MM-6mAPred: Identifying DNA N6-methyladenine sites based on Markov model," *Bioinformatics*, vol. 36, no. 2, pp. 388–392, Jul. 2019.

[20] S. Basith, B. Manavalan, T. H. Shin, and G. Lee, "SDM6A: A Web-based integrative machine-learning framework for predicting 6 mA sites in the rice genome," *Mol. Therapy Nucleic Acids*, vol. 18, pp. 131–141, Dec. 2019.

[21] H. Lv, F.-Y. Dao, Z.-X. Guan, D. Zhang, J.-X. Tan, Y. Zhang, W. Chen, and H. Lin, "IDNA6mA-rice: A computational tool for detecting N6-methyladenine sites in rice," *Frontiers Genet.*, vol. 10, p. 793, Sep. 2019.

[22] H. Yu and Z. Dai, "SNNRice6mA: A deep learning method for predicting DNA N6-methyladenine sites in rice genome," *Frontiers Genet.*, vol. 10, p. 1071, Oct. 2019.

[23] D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3642–3649.

[24] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[25] Z. Zhang, Y. Zhao, X. Liao, W. Shi, K. Li, Q. Zou, and S. Peng, "Deep learning in omics: A survey and guideline," *Briefings Funct. Genomics*, vol. 18, no. 1, pp. 41–57, Feb. 2019.

[26] M. U. Rehman and K. T. Chong, "DNA6mA-MINT: DNA-6mA modification identification neural tool," *Genes*, vol. 11, no. 8, p. 898, Aug. 2020.

[27] Q. Zou, P. Xing, L. Wei, and B. Liu, "Gene2vec: Gene subsequence embedding for prediction of mammalian N6-methyladenosine sites from mRNA," *RNA*, vol. 25, no. 2, pp. 205–218, Feb. 2019.

[28] H. Tayara and K. Chong, "Improved predicting of the sequence specificities of RNA binding proteins by deep learning," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, early access, Mar. 18, 2020, doi: 10.1109/TCBB.2020.2981335.

[29] I. Nazari, M. Tahir, H. Tayara, and K. T. Chong, "IN6-methyl (5-step): Identifying RNA N6-methyladenosine sites using deep learning mode via Chou's 5-step rules and Chou's general PseKNC," *Chemometric Intell. Lab. Syst.*, vol. 193, Oct. 2019, Art. no. 103811.

[30] M. Tahir, H. Tayara, and K. T. Chong, "IPseU-CNN: Identifying RNA pseudouridine sites using convolutional neural networks," *Mol. Therapy Nucleic Acids*, vol. 16, pp. 463–470, Jun. 2019.

[31] Y. Ding, J. Tang, and F. Guo, "Human protein subcellular localization identification via fuzzy model on kernelized neighborhood representation," *Appl. Soft Comput.*, vol. 96, Nov. 2020, Art. no. 106596.

[32] Y. Ding, J. Tang, and F. Guo, "Identification of Drug–Target interactions via dual Laplacian regularized least squares with multiple kernel fusion," *Knowl.-Based Syst.*, vol. 204, Sep. 2020, Art. no. 106254.

[33] H. M. D. Kabir, M. Abdar, S. M. J. Jalali, A. Khosravi, A. F. Atiya, S. Nahavandi, and D. Srinivasan, "SpinalNet: Deep neural network with gradual input," 2020, *arXiv:2007.03347*. [Online]. Available: http://arxiv.org/abs/2007.03347

[34] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol.*, vol. 160, no. 1, p. 106, 1962.

[35] C. Zhou, C. Wang, H. Liu, Q. Zhou, Q. Liu, Y. Guo, T. Peng, J. Song, J. Zhang, L. Chen, Y. Zhao, Z. Zeng, and D.-X. Zhou, "Identification and analysis of adenine N6-methylation sites in the rice genome," *Nature Plants*, vol. 4, no. 8, pp. 554–563, Aug. 2018.

[36] Z. Chen, P. Zhao, F. Li, T. T. Marquez-Lago, A. Leier, J. Revote, Y. Zhu, D. R. Powell, T. Akutsu, G. I. Webb, K.-C. Chou, A. I. Smith, R. J. Daly, J. Li, and J. Song, "ILearn: An integrated platform and meta-learner for feature engineering, machine-learning analysis and modeling of DNA, RNA and protein sequence data," *Briefings Bioinf.*, vol. 21, no. 3, pp. 1047–1057, May 2020.

[37] M. Vogt, "An overview of deep learning and its applications," in *Fahrerassistenzsysteme*. Wiesbaden, Germany: Springer, 2019, pp. 178–202.

[38] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[39] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: http://arxiv.org/abs/1502.03167

[40] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. ICML*, Jan. 2010, pp. 807–814.

[41] M. Vohra, A. Alexanderian, H. Guy, and S. Mahadevan, "Active subspace-based dimension reduction for chemical kinetics applications with epistemic uncertainty," *Combustion Flame*, vol. 204, pp. 152–161, Jun. 2019.

[42] J. Rafferty, P. Shellito, N. H. Hyman, and W. D. Buie, "Practice parameters for sigmoid diverticulitis," *Diseases Colon Rectum*, vol. 49, no. 7, pp. 939–944, Jul. 2006.

[43] R. Caruana, S. Lawrence, and C. L. Giles, "Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 402–408.

[44] B. W. Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme," *Biochimica et Biophysica Acta (BBA) Protein Struct.*, vol. 405, no. 2, pp. 442–451, Oct. 1975.

[45] O. Raymond, J. Gouzy, J. Just, H. Badouin, M. Verdenaud, A. Lemainque, P. Vergne, S. Moja, N. Choisne, C. Pont, and S. Carrere, "The Rosa genome provides new insights into the domestication of modern roses," *Nature Genet.*, vol. 50, no. 6, pp. 772–777, 2018.

[46] P. P. Edger, R. VanBuren, M. Colle, T. J. Poorten, C. M. Wai, C. E. Niederhuth, E. I. Alger, S. Ou, C. B. Acharya, J. Wang, and P. Callow, "Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry *(Fragaria vesca)* with chromosome-scale contiguity," *GigaScience*, vol. 7, no. 2, Feb. 2018, Art. no. gix124.

ZEESHAN ABBAS received the B.Sc. degree in electrical engineering from the COMSATS Institute for Information Technology, Islamabad, Pakistan, in 2011, and the M.Sc. degree in information security from Air University at Islamabad, Islamabad, in 2019. He is currently pursuing the Ph.D. degree in electronics and information engineering from Jeonbuk National University, Jeonju, South Korea. His research interests include artificial intelligence, machine learning, and bio-medical imaging.

HILAL TAYARA received the B.Sc. degree in computer engineering from Aleppo University, Aleppo, Syria, in 2008, and the M.S. and Ph.D. degrees in electronics and information engineering from Jeonbuk National University, Jeonju, South Korea, in 2015 and 2019, respectively. He served as a Researcher with Jeonbuk National University. He is currently serving as an Assistant Professor with the School of International Engineering and Science, Jeonbuk National University. His research interests include bioinformatics, machine learning, and image processing.

KIL TO CHONG (Member, IEEE) received the Ph.D. degree in mechanical engineering from Texas A&M University in 1993. He is currently a Professor with the School of Electronics and Information Engineering, Jeonbuk National University, Jeonju, South Korea, and the Head of the Advanced Information and Electronics Research Center, Jeonbuk National University. He is the President of the Korean Electronics Engineering Society, Systems, and Control. His research interests include the areas of bioinformatics, artificial intelligence, brain disease, and drug discovery.