Lightweight Vision Transformer for Efficient Influenza Virus Subtype Classification via Genomic Image Processing

1st Vatsal Shah
Electrical and Computer Engineering
University of Windsor
Windsor, Canada
shah7r1@uwindsor.ca
*Corresponding author

2nd Love Fadia
Electrical and Computer Engineering
University of Windsor
Windsor, Canada
fadial@uwindsor.ca

3rd Mohammad Hassanzadeh*
Electrical and Computer Engineering
University of Windsor
Windsor, Canada
mhassan@uwindsor.ca

4th Majid Ahmadi
Electrical and Computer Engineering
University of Windsor
Windsor, Canada
M.ahmadi@uwindsor.ca

5th Jonathan Wu
Electrical and Computer Engineering
University of Windsor
Windsor, Canada
jwu@uwindsor.ca

***Abstract*—This study offers a novel method for categorizing influenza virus subtypes. Types A, B, C, and D influenza viruses exhibit different degrees of severity and transmission, making precise classification essential for public health. Using strategies like Gramian Angular Summation Field and Gramian Angular Difference Field and their modified version this technology transforms DNA sequences into visual patterns. Lightweight models such as MobileNetV2 and Vision Transformers with lower number of parameters are then used to process these patterns in order to identify subtypes with a high degree of accuracy. Our method is efficient and less labor-intensive than older methods, and it achieves a 98.34% classification accuracy.**

***Index Terms*—Influenza virus, Genomic Data Analysis, Grammian Angular Summation Field, Convolutional Neural Network and Vision Transformer**

I. Introduction

With different levels of severity and transmissibility, influenza viruses—which are divided into types A, B, C, and D—are important causes of respiratory diseases in both people and animals [1]. For instance, Influenza B is more frequently linked to seasonal epidemics, but Influenza A is known for its capacity to trigger pandemics because of its high mutation rates and zoonotic transmission [2]. On the other hand, Influenza D mainly affects cattle but can also be dangerous to people and influenza C typically results in milder respiratory diseases [3]. For containment, surveillance, and treatment plans to be successful, various influenza types must be accurately classified. Due to advances in recent research, deep learning models can easily able to classify the various different version of this virus, Nowadays, due to the advancement of deep learning approaches several models are being used for robust classification thanks to these image-based representations' ability to capture distinctive viral patterns However, very few of these methods are computationally undemanding since most of the researchers focus on getting better accuracies and precision [4]. We plan to resolve this issue by proposing a lightweight model, which gives better accuracy than other similar sized models in the literature and can be deployed in low computational scenarios. We achieve this by using techniques like Gramian Angular Summation Fields (GASF), Gramian Angular Difference Fields (GADF), Angular Division Grammian Matrix (ADGM), Grammian Difference Tangent Matrix (GDTM) and their modified version the DNA sequences of these viruses are converted into picture representations. After that, these images can be fed to Convolutional Neural Network (CNNs) or Vision Transformers (ViTs) for classification. Here, we have chosen MobileNetV2 a CNN model for comparision with our method that is based on ViT, since they have similar number of parameters, which are around 3.5 million making it ideal for comparison with our model.

II. Related Work

This section examines important research on classification of variants for influenza virus with an emphasis on different machine learning and analytical techniques. Chrysostomou *et.al* developed a predictive model based on Neuraminidase genes that uses signal-processing techniques to detect influenza A virus subtypes (H1N1, H2N2, H3N2, and H5N1) [5]. With a high accuracy of 98.3% through 5-fold cross-validation, it uses a Support Vector Machine (SVM) classifier, Discrete Fourier Transform (DFT) for feature extraction, and F-score for feature selection. This demonstrates its potential for bioinformatics applications in viral subtype classification. on the other hand, In a systematic review, Borkenhagen *et al* examined 49 research papers that used machine learning

algorithms to predict influenza viral characteristics using proteomic or genomic data. The review identifies gaps in model design and makes recommendations for improvements to improve usability in surveillance applications by highlighting techniques such as random forests, SVM, and neural networks for tasks like host discrimination, subtype assignment, and antiviral resistance prediction [6]. Moreover, Marquez *et al* used supervised machine learning techniques for influenza classification based on a clinical dataset of 15,480 records from Mexico City. In areas where RT-qPCR testing may be limited or impossible, their method, which employed Random Forest and Bagging classifiers, demonstrated good accuracy (86%) and specificity (88%), especially helping with diagnosis [7]. Additionally, utilizing information from the Influenza Research Database and Human Surveillance Records, Kumar *et al* created ensemble learning models for early influenza identification. For particular age groups, they achieved up to 91% accuracy by using machine learning approaches including Random Forest, Gradient Boosted Trees, and stacked ensemble algorithms [8]. This study emphasizes the value of ensemble models as effective, faster diagnostic instruments. The table 3 compares the proposed methodology with other state of art methods.It is evident that the proposed methodology, which utilizes Gramian transformations and ViT, achieves the highest accuracy of 98.34% among the compared methods outperforming other state-of-the-art techniques, including ensemble learning and signal processing methods, while leveraging a comprehensive dataset of DNA sequences.

III. Methodology

The steps for each component are described in this section and are illustrated in Fig 1. We started by collecting DNA sequences from the NCBI Virus Database, and gathered 1,000 sequences of varying sequence between 1000-1500 bp each of different genotypes of Influenza. After that we convert these sequences into the images which is based on GASF, GASD and their modified version. Then we apply MobileNetV2 and ViT for classification.

A. Coding Techniques

To construct a grammian matrix, the data must be in numerical format, so EIIP coding was ultilised, which counts the Electron-Ion Potential value of the nucleotide depending on its molecular weight. The nucleotide's four constituents, A, C, G, and T, are transformed into floating point numbers as illustrated below [9].

- Adenine (A): 0.1260
- Guanine (G): 0.0806
- Thymine (T): 0.1335
- Cytosine (C): 0.1340

B. Gramian Angular Summation Field

The GASF is a method for encoding data into a two-dimensional matrix for better visualization and input into the models. It uses trigonometric transformations to map numerical coded DNA sequence into a polar coordinate system and computes the Grammian matrix [10]. The GASF is computed as [10]:

$$G = \cos(\phi_i + \phi_j), \quad \forall i, j \in [1, n], \tag{1}$$

where $\phi_i = \arccos(x_i)$, and x_i is the normalized value of the original data. Each element represents the summation of angles in polar space. This encoding retains the temporal dependency of the original data while providing a structured format suitable for machine learning models.

C. Gramian Angular Difference Field

The GADF is an effective technique for converting sequential data into a two-dimensional matrix that captures temporal or spatial correlations. While GADF is commonly used for numerical time-series data, it may also be used to encode DNA sequences into numerical values [11].

The normalized DNA sequence is mapped to angular values using the arc-cosine function [12]:

$$\phi_i = \arccos(x_i), \quad \phi_i \in [0, \pi]. \tag{2}$$

The GADF matrix encodes the angular differences between nucleotides in the sequence using the sine of their angular differences:

$$\text{GADF}(i, j) = \sin(\phi_i - \phi_j), \tag{3}$$

where ϕ_i and ϕ_j are the angular values corresponding to nucleotides at positions i and j.

D. Angular Division Grammian Matrix

The ADGM transforms data into a 2D matrix by encoding angular differences in polar coordinates. For a normalize data $\{x_t\}_{t=1}^{n}$, the angular values are calculated as $\phi_t = \arccos(x_t)$, where $x_t \in [-1, 1]$. This matrix is then defined as:

$$G_{\text{div}}(i, j) = \cos(\phi_i / \phi_j), \quad \forall i, j \in [1, n]. \tag{4}$$

This symmetric matrix captures temporal dependencies through phase angle differences.

E. Grammian Difference Tangent Matrix

This improved form of the classic Grammian Angular Field (GAF) was created especially to depict numerical coded DNA sequences. By using the tangent of angular differences to capture the summing relationships between the sequence's angular components, GASD provides enhanced sensitivity to minute phase fluctuation.

Each element is now mapped to an angular coordinate using the arcsine function:

$$\theta_t = \arcsin(x_t), \quad x_t \in [-1, 1]. \tag{5}$$

The modified version can be mathematically defined as:

$$G_{\text{sum}}(i, j) = \tan(\theta_i - \theta_j), \quad \forall i, j \in [1, n], \tag{6}$$

where θ_i and θ_j are the angular representations of the i-th and j-th elements of the sequence.

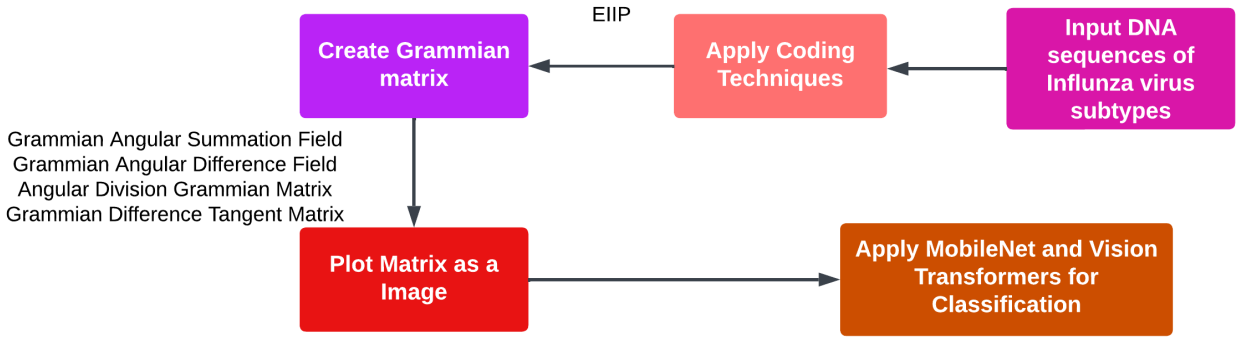

Fig. 1: Methodology for Classification of Different types for Influenza virus

F. MobileNetV2

MobileNetV2 is a CNN architecture specifically designed for mobile and embedded vision applications. It is notable for its ability to balance accuracy and computational cost, making it appropriate for deployment on resource-constrained device3s [13]. MobileNet's main concept is to employ depthwise separable convolutions to greatly reduce the amount of parameters and computations when compared to regular convolutions.

The depthwise separable convolution is expressed as [13]:

$$y = DWConv(x, \{W_{dw}\}) + PWConv(x, \{W_{pw}\}), \quad (7)$$

where $DWConv(x, \{W_{dw}\})$ denotes the depthwise convolution operation, and $PWConv(x, \{W_{pw}\})$ represents the pointwise convolution.

MobileNetV2 consists of the following components, Depthwise separable convolutional layers for efficient feature extraction. Lightweight architecture with reduced parameters for mobile and embedded systems. Global average pooling and a fully connected layer for classification tasks. Configurable hyperparameter (α and ρ) to balance the trade-off between latency and accuracy.

G. Vision Transformer

ViT is a deep learning model that applies the transformer architecture, originally designed for natural language processing, to image recognition tasks [14]. Instead of using convolutional layers, ViT splits an image into fixed-size patches, linearly embeds them, and processes them as a sequence of tokens. It captures long-range dependencies across the image, leveraging self-attention mechanisms to learn global context effectively. ViT has achieved state-of-the-art performance in various computer vision tasks, especially when trained on large datasets, and is known for its scalability and efficiency compared to traditional CNNs. In this implementation, ViT is configured with 6 transformer layers, each with 8 attention heads, using patch sizes of 8x8. The embedding dimensions are set to 512, with a multi-layer perceptron (MLP) size twice the embedding dimension. The model employs learnable positional embeddings, layer normalization, and a classification token (CLS token) for output aggregation. Optimized using Adam with a learning rate of 1e-4 and weight decay of 0.01, this ViT architecture demonstrates adaptability and high performance, making it a robust choice for diverse datasets and tasks.

IV. Results and Discussion

Angle Name	Model Name	Number of Parameters	Testing Accuracy (%)
GADF	Proposed method	3.4 million	98.34
GDTM	Proposed method	3.4 million	98.17
GASF	Proposed method	3.4 million	98.00
ADGM	Proposed method	3.4 million	98.00
GADF	MobileNetV2	3.5 million	92.83
GDTM	MobileNetV2	3.5 million	95.00
GASF	MobileNetV2	3.5 million	94.67
ADGM	MobileNetV2	3.5 million	96.00

TABLE I: Comparison of testing accuracies for different angles and models.

The Table I comparison highlights that the Proposed method outperforms MobileNetV2 across four Genomic Image Processing techniques with consistently higher testing accuracies and fewer parameters approximately (3.4 million vs. 3.5 million). The Proposed method achieves its highest accuracy of 98.33% for the GADF, while accuracies for other Grammian transformation (GASF,ADGM and GDTM) remain close, around 98.00%, showing stability and robustness. In contrast, MobileNetV2 demonstrates more variability, with accuracies ranging from 92.83% for the GADF to 96.00% for the ADGM. Notably, both models achieve their highest performance on the ADGM, though the Proposed method maintains an edge in every scenario. These results underscore the efficiency and consistency of the Proposed method, making it a superior choice for this classification task.

The Table II provides a statistical comparison of accuracy between the Proposed Method and MobileNetV2, analyzing

TABLE II: Statistical Analysis of Accuracy for Proposed Method and MobileNetV2

Type of Network	Mean Accuracy (%)	Std Accuracy (%)	Min Accuracy (%)	Max Accuracy (%)
Proposed Method	98.20	0.25	97.90	98.34
MobileNetV2	94.90	0.35	94.30	95.40

TABLE III: Comparison with State of the Art Methods

Sr No	*Author Name*	*Methodology*	*Dataset*	*Accuracy (%)*
1	Chrysostomou et al.	Signal Processing Techniques	Influenza A subtypes	98.3
2	Marquez .et.al.	Random Forest and Bagging Classifier	15480 dataset of Rt-qPCR tested patient	86
3	Kumar.et.al	Ensemble learning Models	Influenza Research database	91
4	Proposed Methodology	Grammian and its Modified Version + ViT	4000 whole DNA sequences consisting of Influenza A B C D	98.34

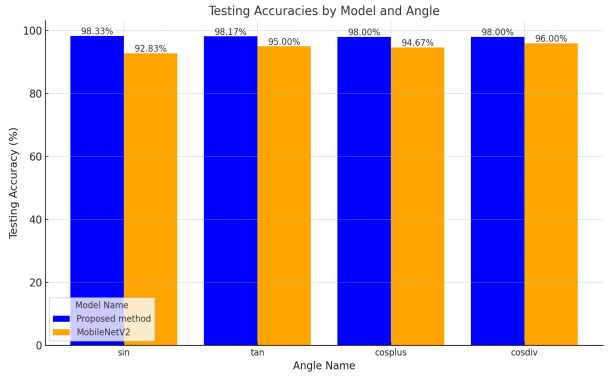

Fig. 2: Comparison of testing accuracy

mean, standard deviation, minimum, and maximum accuracy. The Proposed Method achieves a higher mean accuracy of 98.20%, outperforming MobileNetV2, which has a mean accuracy of 94.90%. Additionally, the Proposed Method demonstrates lower variability with a standard deviation of 0.25%, compared to 0.35% for MobileNetV2, indicating more consistent performance. The accuracy range for the Proposed Method spans 97.90% to 98.50%, whereas MobileNetV2 ranges between 94.30% and 95.40%. These results highlight the superior performance and reliability of the Proposed Method, which achieves better stability and consistently high accuracies across runs. In contrast, MobileNetV2, while competitive, demonstrates greater variability and lower peak performance. Overall, the Proposed Method stands out as the more robust and efficient model for this task.

V. Conclusion and Future Work

When deep learning models like MobileNetV2 and ViTs are combined with genomic image processing, the study indicates that influenza subtype classification accuracy and efficiency significantly increase. The results show that the proposed methods outperform popular ones, offering a dependable and scalable alternative for genetic data analysis. Apart from reducing the preprocessing requirements, the framework demonstrates how similar methods could be used to other genomic classification issues. Future research should examine the applicability of these techniques to bigger datasets and enhance model designs for better performance and generalizability.

References

[1] T. M. Uyeki, D. S. Hui, M. Zambon, D. E. Wentworth, and A. S. Monto, "Influenza," *The Lancet*, vol. 400, no. 10353, pp. 693–706, Aug. 2022.

[2] D. Panatto, A. Signori, P. L. Lai, R. Gasparini, and D. Amicizia, "Heterogeneous estimates of influenza virus types A and B in the elderly: Results of a meta-regression analysis,"*Influenza and Other Respiratory Viruses*, vol. 12, no. 4, pp. 533–543, Mar. 2018, doi: 10.1111/irv.12550.

[3] T. Wolff and M. Veit, "Influenza B, C and D viruses (Orthomyxoviridae)," in *Elsevier eBooks*, 2020, pp. 561–574. doi: 10.1016/b978-0-12-809633-8.21505-7.

[4] Y. Xu and D. Wojtczak, "Dive into machine learning algorithms for influenza virus host prediction with hemagglutinin sequences," *BioSystems*, vol. 220, p. 104740, Oct. 2022, doi: 10.1016/j.biosystems.2022.104740.

[5] C. Chrysostomou and H. Seker, "Signal-processing-based bioinformatics approach for the identification of influenza A virus subtypes in neuraminidase genes," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2013, pp. 3066–3069, doi: 10.1109/EMBC.2013.6610188.

[6] L. K. Borkenhagen, M. W. Allen, and J. A. Runstadler, "Influenza virus genotype to phenotype predictions through machine learning: a systematic review," *Emerging Microbes Infections*, vol. 10, no. 1, pp. 1896–1907, Jan. 2021, doi: 10.1080/22221751.2021.1978824.

[7] E. Marquez, E. V. Barrón-Palma, K. Rodríguez, J. Savage, and A. L. Sanchez-Sandoval, "Supervised Machine Learning Methods for Seasonal Influenza Diagnosis," *Diagnostics (Basel)*, vol. 13, no. 21, p. 3352, Oct. 2023, doi: 10.3390/diagnostics13213352.

[8] R. Kumar, S. Maheshwari, A. Sharma, S. Linda, S. Kumar, and I. Chatterjee, "Ensemble learning-based early detection of influenza disease," *Multimed Tools Appl.*, pp. 1–21, May 2023, doi: 10.1007/s11042-023-15848-2.

[9] A. S. Nair and S. P. Sreenadhan, "A coding measure scheme employing electron-ion interaction pseudopotential (EIIP)," *Bioinformation*, vol. 1, no. 6, pp. 197–202, Oct. 2006.

[10] C. L. Yang, C. Y. Yang, Z. X. Chen and N. W. Lo, "Multivariate Time Series Data Transformation for Convolutional Neural Network," *2019 IEEE/SICE International Symposium on System Integration (SII)*, Paris, France, 2019, pp. 188-192, doi: 10.1109/SII.2019.8700425.

[11] M. Lin, J. Wu, J. Meng, W. Wang, and J. Wu, "Screening of retired batteries with gramian angular difference fields and ConvNeXt," *Engineering Applications of Artificial Intelligence*, vol. 123, no. Part B, p. 106397, Aug. 2023, doi: 10.1016/j.engappai.2023.106397.

[12] Y. Liu, S. Dou, Y. Du, and Z. Wang, "Gearbox Fault Diagnosis Based on Gramian Angular Field and CSKD-ResNeXt," *Electronics*, vol. 12, no. 11, p. 2475, May 2023, doi: 10.3390/electronics12112475.

[13] D. Sinha and M. El-Sharkawy, "Thin MobileNet: An Enhanced MobileNet Architecture," *2019 IEEE 10th Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, New York, NY, USA, 2019, pp. 0280-0285, doi: 10.1109/UEMCON47517.2019.8993089.

[14] W. Dou, K. Wang and T. Yamauchi, "Face Expression Recognition With Vision Transformer and Local Mutual Information Maximization," in *IEEE Access*, vol. 12, pp. 169263-169276, 2024, doi: 10.1109/ACCESS.2024.3496506.