applied sciences

MDPI

Article

Comparison of Monkeypox and Wart DNA Sequences with Deep Learning Model

Talha Burak Alakus [1,*] and Muhammet Baykara [2]

[1] Department of Software Engineering, Kırklareli University, Kırklareli 39000, Turkey
[2] Department of Software Engineering, Fırat University, Elazıg 23119, Turkey
* Correspondence: talhaburakalakus@klu.edu.tr

Abstract: After the COVID-19 disease, monkeypox disease has emerged today and has started to be seen almost everywhere in the world in a short time. Monkeypox causes symptoms such as fever, chills, and headache in people. In addition, rashes are seen on the skin and lumps are formed. Early diagnosis and treatment of monkeypox, which is a contagious disease, are of great importance. An expert interpretation and clinical examination are usually needed to detect monkeypox. This may cause the treatment process to be slow. Furthermore, monkeypox is sometimes confused with warts. This leads to incorrect diagnosis and treatment. Because of these disadvantages, in this study, the DNA sequences of HPV causing warts and MPV causing monkeypox were analyzed and the classification of these sequences was performed with a deep learning algorithm. The study consisted of four stages. In the first stage, DNA sequences of viruses that cause warts and monkeypox were obtained. In the second stage, these sequences were mapped using various DNA-mapping methods. In the third stage, the mapped sequences were classified using a deep learning algorithm. At the last stage, the performances of DNA-mapping methods were compared by calculating accuracy and F1-score. At the end of the study, an average accuracy of 96.08% and an F1-score of 99.83% were obtained. These results showed that these two diseases can be effectively classified according to their DNA sequences.

Keywords: monkeypox; deep learning; DNA mapping; genome converting

check for updates

Citation: Alakus, T.B.; Baykara, M. Comparison of Monkeypox and Wart DNA Sequences with Deep Learning Model. *Appl. Sci.* **2022**, *12*, 10216. https://doi.org/10.3390/app122010216

Academic Editors: Yujin Lim and Hideyuki Takahashi

Received: 12 September 2022
Accepted: 9 October 2022
Published: 11 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Monkeypox is an illness and a viral zoonotic infection caused by the monkeypox virus. It can spread from animals to humans and spread from person to person. It was first described as a zoonosis in endemic areas following the eradication of smallpox in the year 1980. Monkeypox virus is seen sporadically in the rainforest regions of Central and West Africa, particularly in the Democratic Republic of the Congo. Clinically, the disease is indistinguishable from human smallpox, chickenpox, and warts. Unlike other animal pox viruses, the monkeypox virus causes general infections in humans [1]. Monkeypox is clinically manifested by symptoms such as fever, malaise, fatigue, headache, muscle aches, back pain, low energy, rash, and swollen lymph nodes and can cause a range of medical complications. Monkeypox virus has an incubation period of 5 to 21 days, and the febrile stage usually lasts for 1 to 3 days [2].

Monkeypox was first identified in monkeys in laboratory studies in 1958, and that is where its name comes from. However, monkeys are not natural reservoirs. It was first seen in humans in 1970 in the Congo, where the smallpox virus was eradicated in 1968. Many cases of monkeypox encountered since this date have been seen in rural and rainforest areas. Since 1970, in eleven African countries (Benin, Nigeria, Cameroon, Democratic Republic of the Congo, Central African Republic, Gabon, Liberia, Ivory Coast, Republic of the Congo, Sierra Leone, and South Sudan), monkeypox virus has been found in humans. For the world, which is more susceptible to epidemics, especially after the COVID-19 pandemic,

monkeypox can be considered a disease of global importance, as it affects not only West and Central African countries but also the rest of the world, albeit rarely and in small numbers. The first monkeypox epidemic outside of Africa was seen in the United States in 2003. This outbreak has resulted in over seventy cases of monkeypox in the USA. Then, monkeypox occurred in those who traveled from Nigeria to Israel and the United Kingdom in September 2018; to Singapore in May 2019, December 2019, May 2021, and May 2022; and again from Nigeria to the USA in July and November 2021. In May 2022, there were multiple cases of monkeypox in several nonendemic countries. Again, in May 2022, cases were reported in Canada, Australia, Israel, and the United Arab Emirates.

The world, which has been worried about the COVID-19 pandemic for about three years, started to worry about the monkeypox virus this time after the announcements of the World Health Organization. Many local and international research and survey studies have been conducted about monkeypox. Some of the studies provide brief information about monkeypox, which has attracted attention again after the COVID-19 pandemic. These studies generally aim to provide information about the clinical course, epidemiology, diagnosis, treatment, and prevention methods of the disease. While some of the studies [3] have questioned whether there is a need for concern, some have focused on surveys to reveal that monkeypox causes less worry [4]. With similar logic, another study presents whether there is a potential crisis or not, from a general point of view [5].

Since monkeypox causes rashes or pustules on the body, in some cases, these lesions are confused with acne, syphilis, herpes, or warts [6]. An expert opinion is required to distinguish between them. However, in some cases, the difference is not completely clear and causes the diagnosis to be made incorrectly. This situation causes the treatment to be wrong and causes the patient to lose time. To prevent such problems, the need for computer-aided systems has arisen [7]. In computer-aided diagnostic systems, skin images are generally used, and it can be determined whether the disease is monkeypox or another disease on these images. However, expert interpretation is required for images to be labelled. This causes the analysis process to be long. For these reasons, researchers turn to alternative computer-based approaches with less error rate and prefer bioinformatics studies for this. The aim of this study is to suggest using a different method from these methods to avoid the confusions mentioned, and to apply bioinformatics approaches for this. In this way, confusion of the new monkeypox disease with warts will be prevented, the physicians' job will be easier, and the patients will not lose time in terms of treatment and diagnosis. For this, DNA sequences were used in the study and the distinction of these two diseases was made with an artificial intelligence technique. Today, the importance of studies based on bioinformatics and genomic signal processing has increased and it has been used effectively in health fields [8,9]. Since monkeypox is a new disease, although there are not many studies in the field of bioinformatics, certain studies are available in the literature. The study [10] focused on the phylogenomic characterization and microevolutionary manifestations of the multi-country monkeypox virus outbreak. In the study, information about the clade and the origin information of the epidemic genome sequences were obtained. In study [11], a rapid detection method was developed for monkeypox using a recombinase polymerase amplification assay. The researchers reported that the specificity of their method was 100%, while the sensitivity was 95%. In the study [12], the researchers aimed to make a comparison of clinical data via qPCR using oropharyngeal swabs, lesion swabs, and blood. So, this study showed the reliability of cutaneous lesion samples of swabs with the observation of DNA for the detection of monkeypox, which is considered the gold standard for diagnostics. Some studies [13–15] related to the subject have focused on the importance of vaccines and various vaccine studies, especially with the effect of the COVID-19 epidemic. Two vaccines for monkeypox have been approved in the United States. The names of these vaccines are JYENNEOS and ACAM2000. Vaccines are licensed by the FDA. The vaccines are of the live-virus-containing type, and the live virus form in the JYENNEOS vaccine is administered as an injection 28 days apart. The ACAM2000 vaccine, on the other hand, is a percutaneous injection, and it takes four weeks to obtain the result. It

is stated that the vaccines provide benefits for the monkeypox virus, unlike the COVID-19 pandemic [13]. Studies have generally recommended that vaccines should be avoided during pregnancy. Vaccination is recommended in the current outbreak, especially for those who have encountered confirmed case owners, including healthcare workers [14]. In some studies, in the literature, it has been revealed that the price of the vaccine is not a significant barrier to vaccination among doctors. Study [15] provides an Indonesian example of such a study. According to the information we have obtained from the research [16], the structure of most of the proteins of the monkeypox virus is not yet known. It has been found that the AlphaFold2 method is frequently used in the literature to obtain protein structures of virus proteomes. In this study, the protein structures of the proteome of the reference monkey flower were predicted with the AlphaFold2 method and 186 high-fidelity protein structures were obtained.

To find solutions to the aforementioned problems, the BiLSTM (bidirectional long/short term memory) model, which is one of the deep learning models, was used in this study, and the DNA of viruses that cause monkeypox and warts were classified. The study consisted of four different stages. In the first stage, DNA sequences were obtained. In the second stage, DNA sequences were mapped using various DNA-mapping methods. These methods are atomic number, EIIP (electron ion interaction potential), integer number, real number, and molecular mass methods. Numerical expressions of DNA sequences were obtained through these methods and made ready for classification. In the third stage, the deep learning model was designed, and the DNA sequences were classified. At the last stage, the performances of DNA-mapping methods were determined using accuracy, recall, precision, and F1-score evaluation criteria. The highlights of the study are:

- To the best of our knowledge, for the first time in this study, the DNA sequences of viruses that cause warts and monkeypox were analyzed using a DNA-mapping methods and classified with deep learning algorithm.
- With this study, it was observed that the approach based on DNA sequencing may be more effective than visual inspection.

The main motivation of the study is the inadequacy of the visual diagnosis procedures and, accordingly, the wrong treatment. Therefore, computer-aided systems are needed, and automatic diagnosis is required. Monkeypox disease can be confused with warts on the skin in some cases. A certain expert opinion is needed, and the expert should carry out the examination in detail. To prevent this problem, skin images are used, and computer-aided systems are developed. However, expert interpretation is required for images to be labelled. This causes the analysis process to be long. For these reasons, in this study, a healthier and more effective method was preferred, and the distinction of diseases was made according to DNA sequences. Because the DNA of the two diseases is different, in this case, there is no chance of confusing the diagnosis of the two diseases based on DNA sequences. The contributions of the study can be expressed as follows:

- In this study, the distinction between monkeypox and warts was made according to the DNA sequences and the similarity between these two diseases was eliminated. There is no confusion as the DNA sequence is different in the two diseases.
- Since the distinction between the two diseases is based on DNA sequences, a visual inspection is out of the question. In this way, specialists will not have direct contact with patients and will not examine them manually.
- By obtaining the DNA sequence of the virus seen in the patient, the diagnosis of the disease can be made easily. In this way, both warts and monkeypox will not be confused and there will be no wrong treatment.

The organization of the study is as follows. In Section 2, information about the data set and deep learning model used in the study is given. In addition to these, DNA-mapping methods are also mentioned. In Section 3, the results of the application were examined and discussed. In addition, the advantages and disadvantages of the study are also emphasized.

In Section 4, the study is summarized and the highlights of the study are given and future studies are mentioned.

2. Material and Methods

2.1. MPV and HPV Data Set

In this study, DNA sequences of MPV (Monkeypox Virus) and HPV (Human Papilloma Virus) virus were used, and a distinction between these two viruses was made by classifying them with these sequences. A total of 110 genome sequences, 55 each, were used for the MPV and HPV virus. The reason the genome count is so low is that MPV is a new virus and does not have as many data as HPV. Therefore, the data set is unbalanced. In order to avoid this problem, the zero-padding method, which is one of the methods frequently used in bioinformatics fields, was used [17]. To eliminate the imbalance between these data, the lowest number of viruses was taken as a basis and the data were obtained with the same number. The highest DNA sequence length for HPV virus is 7904 while the lowest DNA sequence length is 409. The highest DNA sequence length used for MPV was 198,740, while the lowest DNA sequence length was 942. According to the information given, the DNA sequence lengths for both virus types were not equal. With this method, 0 was added to the end of the DNA sequences and this process was performed until the highest DNA sequence length was obtained. Since the maximum sequence length for this study was 198,740, 0 was added to the end of all DNA sequences until the lengths of all DNAs were 198,740. In this way, all data had equal length and the imbalance was eliminated. Table 1 contains information showing the final state of the data.

Table 1. Information about the data set in the study.

Virus Type	Num. of Genoms	Sequence Length	Num. of Features
MPV	55	198,740	10,930,700
HPV	55	198,740	10,930,700

After the data set was created, all DNA sequences were converted into numerical expressions with various DNA-mapping techniques and classified.

2.2. DNA-Mapping Methods

In order for DNA sequences to be evaluated using artificial intelligence methods, the sequences must be converted into numerical expressions. DNA sequences consist of four bases. These bases are adenine (A), thymine (T), cytosine (C) and guanine (G). There are methods in the literature that can convert these bases into various numerical expressions. In this study, some of these methods were used and DNA sequences were mapped. In line with the study, five different methods were used: integer number representation, real number representation, atomic number representation, EIIP, and molecular mass. The integer-mapping method is widely used in DNA studies [18]. In this method, DNA bases are first listed alphabetically and after the sorting process, integer values are assigned to the bases. In this direction, the value is 1 for the A base, 2 for the C base, 3 for the G base, and 4 for the T base. For example, a DNA sequence S(n) = [ACTGCTAGC] is mapped as C(n) = [124324132] with the integer number representation method. In the real-number-mapping technique, the bases are assigned values of −1.5 and 0.5. These values differ for bases. In this method, the A base takes the value −1.5, while the T base takes the value 1.5. In addition, the C base is expressed with the value of 0.5, while the G base is converted into a numerical expression with the value −0.5. The integer and real-number-mapping methods are frequently used in studies carried out with artificial intelligence. The biggest reason for this situation is that the deviations in these methods are symmetrical [19,20]. For example, a DNA sequence of S(n) = [ACTGCTAGC] is mapped as C(n) = [−1.5 0.5 1.5 −0.5 0.5 1.5 −1.5 −0.5 0.5 using the real mapping method. In the molecular mass DNA-mapping method, the numerical values of the bases are determined

based on their molecular masses [21,22]. In this method, A = 134, C = 110, G = 150, and T = 125 values are used. For example, a DNA sequence S(n) = [ACTGCTAGC] is converted into a numerical expression such as C(n) = [134 110 125 150 110 125 134 150 110] by the molecular mass DNA-mapping method. Integer number, real number, and molecular mass mapping methods are evaluated in the category of fixed mapping methods. The atomic number and EIIP methods are the leading physicochemical-based mapping methods. The atomic numbers of the bases are considered in the atomic mapping method [23]. In line with this method, the A base takes the value 70 and the G base takes the value 78. Moreover, the C base is denoted with the value 58, while the T base is expressed by the value 66. For example, a DNA sequence given as S(n) = [ACTGCTAGC] is converted into numerical expressions such as C(n) = [70 58 66 78 58 66 70 78 58] by the atomic number mapping method. In the EIIP DNA-mapping method, each nucleotide found in the sequence is matched with the half-valence numbers in the EIIP representation [24]. In this method, the A base takes the value 0.1260, the G base takes the value 0.0806, the C base takes the value 0.1340, and the T base takes the value 0.1335. For example, a DNA sequence given as S(n) = [ACTGCTAGC] is converted into numerical expressions such as C(n) = [0.1260 0.1340 0.1335 0.0806 0.1340 0.1335 0.1260 0.0806 0.1340] by the EIIP DNA-mapping method.

2.3. Deep Learning Model

Deep learning, which is widely used in many fields such as social media, finance, health services, cyber security, and digital assistantship, is defined as a more complex form of artificial neural networks that work like the human brain in its simplest form. With deep learning, using very large data sets, high-level new data are derived without the need for external intervention or human factors, by completing the processes of learning from low-level features, memorizing, and revealing the relationship between data. In this context, deep learning is closely related to artificial intelligence, which has become very popular today, and is closely related to machine learning. A memory is added to the artificial neural network with RNNs (recurrent neural networks), which is one of the widely used deep learning methods. Thus, the neural network will produce an output considering the inputs it has received before. There is a self-feeding loop in the RNN, where many neurons in the network are evaluated. Since neurons connect with the previous neuron over time, the currently working neural network can receive information from the previous neuron. RNN creates networks with loops to make the information permanent. After these networks are created, the output of each layer is given as input to the next hidden layer. Thus, every previous output is learned. If the traditional RNN working in this way has only a short-term memory, it causes problems such as gradient descent when the depth of neural networks increases [25]. One of the RNN architectures, BiLSTM, is a neural network used in natural language processing. The input given in the BiLSTM architecture flows in two directions, from right to left and from left to right. Thus, another LSTM layer is added to the LSTM layer in the opposite direction. In this context, a strong structure is obtained by modeling the sequential dependencies between words and expressions in both directions of the input. BiLSTM consists of two forward and backward LSTM networks. In this way, BiLSTM provides additional training to the data by crossing the data given as an input in both directions. With the forward LSTM layer, chronological data are considered. The forward layer is important for prediction. The backward LSTM layer protects the previous and next information of the system. BiLSTM provides recognition by memorizing long data series [26]. Due to this structure, the BiLSTM architecture produces efficient results in various fields such as sentence classification, speech recognition, and natural language processing. Due to these advantages of the BiLSTM model, this model structure was used in this study and DNA sequences were classified. The flow chart of the study is given in Figure 1.

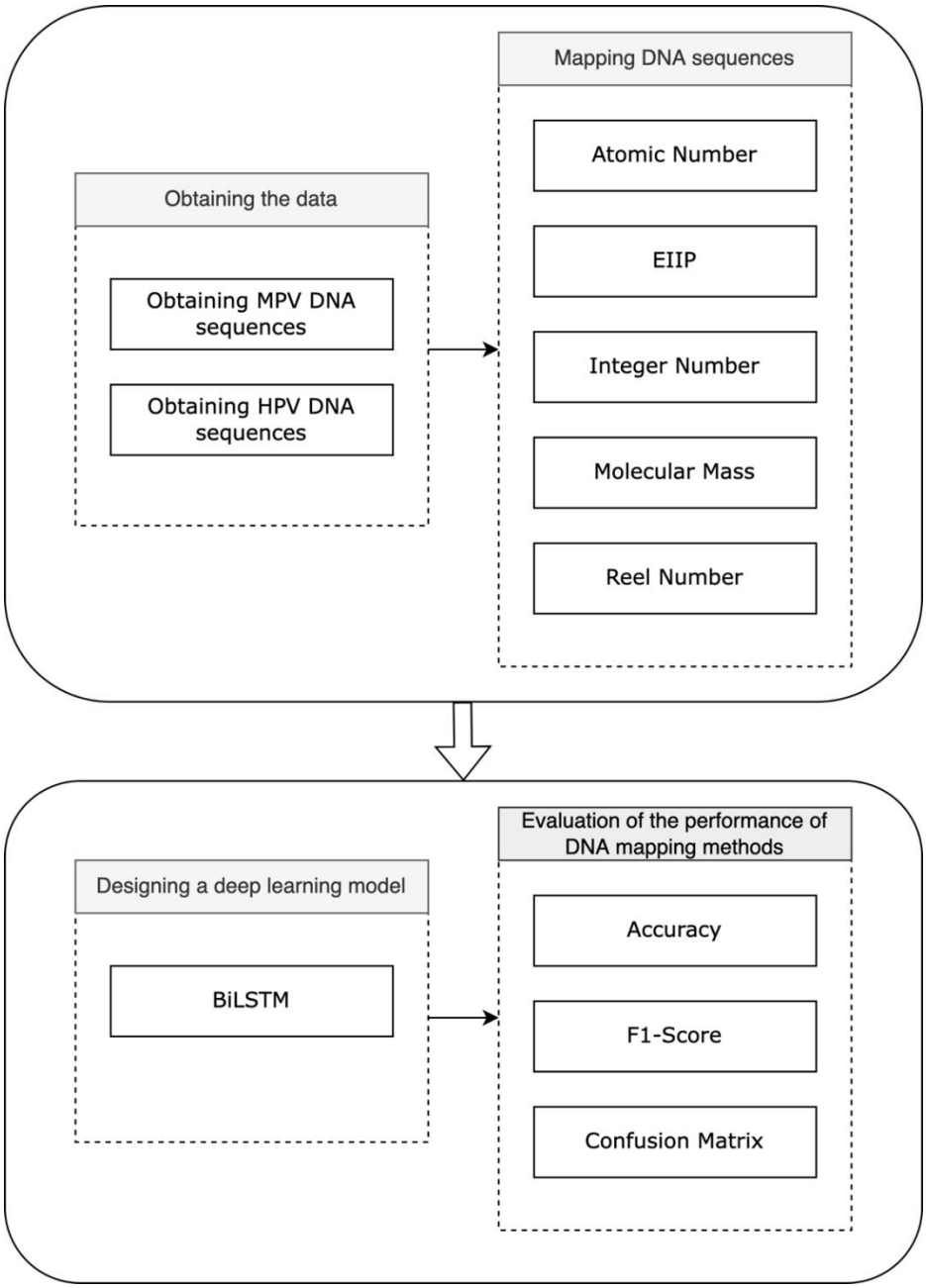

Figure 1. Flowchart of the study.

3. Application Results and Discussion

In this study, DNA sequences of MPV and HPV viruses were used and classified. In the study, DNA sequences were converted into numerical expressions with various DNA numerical mapping methods and the performances of these methods were evaluated with accuracy, F1-score, precision, and recall evaluation metrics. The parameters of the BiLSTM model used in the study were determined by a trial-and-error approach and the parameters producing the most successful results were considered. The parameters of the deep learning model are given in Table 2. While performing the classification with the model, 80% of the data were trained and the remaining 20% were used for testing. This approach was preferred due to the large data set size. The results of the classification process are given in Table 3.

Table 2. Parameters of the designed deep learning model.

Type of Parameters	Value
Number of LSTM units	64
Number of layers	1
Activation function	ReLU (Rectified Linear Unit)
Learning rate	0.0001
Loss function	Categorical crossentropy
Number of epochs	500
Optimization	Adam
Decay	0.000001
Momentum	0.3
Number of fully connected layers	2
Number of fully connected neurons	256 and 128, respectively
Dropout rate	0.20

Table 3. Classification performances of DNA-mapping methods.

DNA-Mapping Methods	Accuracy Rate	Precision Rate	Recall Rate	F1-Score
Atomic number	95.57%	100%	99.44%	99.71%
EIIP	97.65%	99.92%	100%	99.95%
Integer number	99.50%	100%	99.53%	99.76%
Molecular mass	95.84%	100%	99.53%	99.76%
Reel number	91.85%	99.91%	100%	99.95%

According to the classification results given in Table 3, all DNA-mapping methods performed an effective classification process. The lowest classification process was obtained with the real number DNA-mapping method, and an accuracy score of 91.85% was obtained with this method. All the remaining DNA-mapping methods showed an accuracy of over 95%. Among these methods, the lowest accuracy value was obtained with atomic number DNA-mapping method and molecular mass DNA-mapping method, and the accuracy scores of these methods were 95.57% and 95.84%, respectively. The most effective classification was carried out by the EIIP DNA-mapping and integer number DNA-mapping methods. While 97.65% accuracy was achieved with the EIIP DNA-mapping method, this rate increased to 99.50% in the integer number DNA-mapping method. In cases where the data set is unbalanced, the accuracy score alone may not be an adequate evaluation criterion [27]. Although the data set used in the study was balanced, the performances of other evaluation criteria were also calculated, and the performances of DNA-mapping methods were also evaluated with these criteria. As the values of recall, precision, and F1-score get closer to 100%, the performance of the classifier becomes more effective. With this study, the success of all DNA-mapping methods was over 99%. This proved that the methods were effective in the classification process. The confusion matrix of all DNA-mapping methods is given in Figure 2.

Since the new monkeypox disease is still in its early stages, a large number of data are not available. However, the designed deep learning model effectively classified MPV and HPV, with an average accuracy of 96.08%. Although the results were high, each DNA-mapping method produced different results, as seen in Table 3. The lowest accuracy score was obtained from the real number, atomic number, and molecular mass DNA-mapping methods. The real number and molecular mass DNA-mapping methods are fixed mapping methods. In short, there is no need for certain knowledge (structure of bases, chemical properties, etc.) while performing the mapping process with these methods. In the real number DNA-mapping method, numbers are assigned to the bases in the DNA sequences, while the masses of the DNA bases are used in the molecular mass DNA-mapping method. This may have caused the information in the DNA sequences to be lost. The atomic number DNA-mapping method, on the other hand, is a physicochemical-based method, unlike these two methods. In this method, the mapping process was carried out according to the

chemical properties of the bases. The reason why this method was less ineffective than real number and molecular mass DNA-mapping methods may be that the chemical properties of bases were not used in the study. The use of chemical properties may increase the performance of this method. The most effective classification process was carried out with the EIIP and integer number DNA-mapping methods. The EIIP DNA-mapping method is a physicochemical-based method just like the atomic number DNA-mapping method. The reason why this method was successful over the atomic number DNA-mapping method may be that its size takes up less space. In artificial intelligence applications, the size of the data is of great importance and affects the classification performance. In addition, the use of chemical properties of bases may positively affect the performance of this method. The most effective classification result was obtained by the integer number DNA-mapping method. In this method, the mapping process is not based on specific information. However, it is a method often used in artificial intelligence applications. The main reason for this is that the deviations are symmetrical in this method [19]. This has a positive effect on the performance of the classifier.

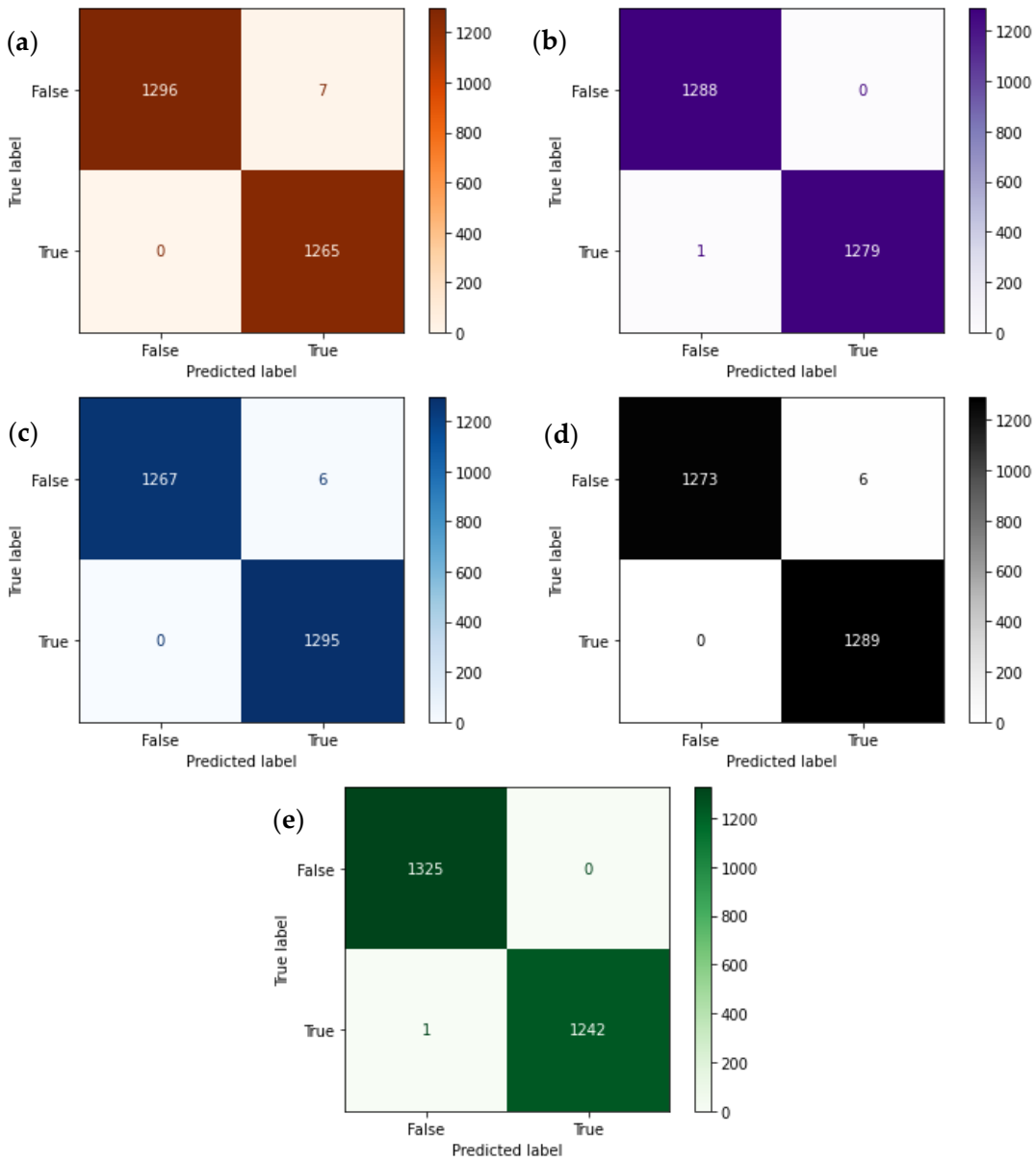

Figure 2. Confusion matrix results of DNA-mapping methods. (**a**) Atomic number, (**b**) EIIP, (**c**) Integer number, (**d**) Molecular mass, (**e**) Real number.

The advantages and disadvantages of the study can be expressed as follows:

- The number of data is of great importance in deep learning studies. Although the number of data used in this study was small, an effective classification process was carried out. In the future, the DNA sequences of the virus will multiply, and more data will be obtained. Since the results of the classification process with more data will be healthier, the results of the findings obtained in this study may vary.
- Furthermore, only one of the deep learning models, the BiLSTM model, was used in this study. The use of different deep learning models or the use of machine learning algorithms will pave the way for the comparison of the study in terms of the literature and will show the effect of computer-aided approaches.
- Finally, it is sometimes difficult to determine whether human bumps are monkeypox or warts. Even experienced specialists cannot distinguish them clearly in some cases, and this may cause the diagnosis and treatment to be wrong. Diagnosis based on DNA sequences, which was the starting point of this study, was more effective. In this way, the distinction between these two diseases can be made clearly. Obtaining the DNA sequence of the disease seen in the body and analyzing this sequence with computer-aided applications will be healthier in terms of both diagnosis and confidence. The application results obtained support this.
- The performance of the model varied according to the DNA-mapping methods used. This reveals the biggest limitation of this study. The absence of a standard method and the variation in the model according to the mapping methods reduce the applicability of the model. It takes time to identify and implement an appropriate mapping method.
- In this study, raw DNA sequences were used and classified only by mapping methods. Performance can be increased by applying various feature extraction operations (signal processing, image processing, etc.).
- In addition, a secure and distributed environment can be created to share data. In this way, the number of data can be increased and more robust and reliable applications can be developed. Blockchain technology, one of the new technologies, can be applied to this field and evaluated in future studies. It has been observed that similar applications are made for COVID-19 disease, which has become a pandemic today [28].
- Furthermore, some optimization algorithms can be used, and these results can be improved. Optimization algorithms are frequently used in data classification studies and can positively affect the classification result [29]. It is important to use optimization algorithms in future studies.
- In this study, DNA sequences were used and new monkeypox and wart diseases were predicted from the DNA sequences. However, diseases can be predicted with RNA sequences and effective results are obtained [30]. In the future, RNA sequences of new monkeypox and wart disease should be used and the results obtained in this study should be supported.

4. Conclusions

In this study, DNA sequences of viruses that cause monkeypox and warts were used, and a distinction was made between these two diseases by classifying the sequences. The study consisted of four different stages. In the first stage, DNA sequences of MPV and HPV were obtained. In the second stage, DNA sequences were mapped with five different methods, namely integer number, atomic number, EIIP, molecular mass, and real number DNA-mapping methods. In the third stage, a deep learning model was designed and BiLSTM was used in this direction. At the last stage, classification was made and the performances of DNA-mapping methods were determined using accuracy, recall, precision, and F1-score evaluation criteria. Five different DNA-mapping methods used made an effective classification and an average accuracy of 96.08% was obtained. At the end of the classification stage, the most ineffective classification result was obtained with the real number DNA-mapping method and an accuracy score of 91.85% was calculated. The accuracy scores of atomic number and molecular mass DNA-mapping methods were

close to each other. While an accuracy score of 95.84% was obtained in the molecular mass method, this score was 95.57% in the atomic number DNA-mapping method. The most effective accuracy values were reached by EIIP and integer number DNA-mapping methods. These two methods showed an accuracy of over 97%. While 97.65% accuracy was obtained with the EIIP method, 99.5% accuracy was observed with the integer number DNA-mapping method. In line with the findings obtained as a result of the study, it was observed that the classification and DNA-mapping methods used were effective in determining wart and monkeypox disease. In future studies, different deep learning or machine learning algorithms will be used and the findings obtained in this study will be supported. With this study, a study in this field has been added to the literature and it is hoped that this study will lead to future studies.

Author Contributions: Conceptualization, T.B.A.; methodology, T.B.A. and M.B.; software, T.B.A. and M.B.; writing—original draft preparation, T.B.A. and M.B.; writing—review and editing, T.B.A. and M.B.; funding acquisition, M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All information about the data used in this study are explained in Section 3.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. Monkeypox. Available online: https://www.who.int/news-room/fact-sheets/detail/monkeypox (accessed on 8 September 2022).
2. World Health Organization. Monkeypox. Available online: https://www.who.int/health-topics/monkeypox#tab=tab_2 (accessed on 8 September 2022).
3. Shaheen, N.; Diab, R.A.; Meshref, M.; Shaheen, A.; Ramadan, A.; Shoib, S. Is there a need to be worried about the new monkeypox virus outbreak? A brief review on the monkeypox outbreak. *Ann. Med. Surg.* **2022**, *81*, 104396. [CrossRef] [PubMed]
4. Temsah, M.H.; Aljamaan, F.; Alenezi, S.; Alhasan, K.; Saddik, B.; Al-Barag, A.; Alhaboob, A.; Bahabri, N.; Alshahrani, F.; Alrabiaah, A.; et al. Monkeypox caused less worry than COVID-19 among the general population during the first month of the WHO Monkeypox alert: Experience from Saudi Arabia. *Travel Med. Infect. Dis.* **2022**, *49*, 102426. [CrossRef] [PubMed]
5. Ranganath, N.; Tosh, P.K.; O'horo, J.; Sampathkumar, P.; Binnicker, M.J.; Shah, A.S. Monkeypox 2022: Gearing up for Another Potential Public Health Crisis. *Mayo Clin. Proc.* **2022**, *97*, 1694–1699. [CrossRef] [PubMed]
6. Centers for Disease Control and Prevention. Intervention Services for People with or Exposed to Monkeypox. Available online: https://www.cdc.gov/poxvirus/monkeypox/health-departments/intervention-services.html (accessed on 8 September 2022).
7. Ali, S.N.; Ahmed, T.; Paul, J.; Jahan, T.; Sani, S.M.S.; Noor, N.; Hasan, T. Monkeypox Skin Lesion Detection Using Deep Learning Models: A Feasibility Study. *arXiv* **2022**, arXiv:2207.03342.
8. Naeem, S.M.; Mabrouk, M.S.; Marzouk, S.Y.; Eldosoky, M.A. A diagnostic genomic signal processing (GSP)-based system for automatic feature analysis and detection of COVID-19. *Brief. Bioinform.* **2021**, *22*, 1197–1205. [CrossRef]
9. Naeem, S.M.; Mobrouk, M.S.; Eldosoky, M.A.; Sayed, A.Y. Automated detection of colon cancer using genomic signal processing. *Egypt. J. Med. Hum. Genet.* **2021**, *22*, 77. [CrossRef]
10. Isidro, J.; Borges, V.; Pinto, M.; Sobral, D.; Santos, J.D.; Nunes, A.; Mixão, V.; Ferreira, R.; Santos, D.; Duarte, S.; et al. Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. *Nat. Med.* **2022**, *28*, 1569–1572. [CrossRef]
11. Davi, S.D.; Kissenkötter, J.; Faye, M.; Böhlken-Fascher, S.; Stahl-Hennig, C.; Faye, O.; Faye, O.; Sall, A.A.; Weidmann, M.; Ademowo, O.G.; et al. Recombinase polymerase amplification assay for rapid detection of Monkeypox virus. *Diagn. Microbiol. Infect. Dis.* **2019**, *95*, 41–45. [CrossRef]
12. Nörz, D.; Brehm, T.T.; Tang, H.T.; Grewe, I.; Hermanussen, L.; Matthews, H.; Pestel, J.; Degen, O.; Günther, T.; Grundhoff, A.; et al. Clinical characteristics and comparison of longitudinal qPCR results from different specimen types in a cohort of ambulatory and hospitalized patients infected with monkeypox virus. *J. Clin. Virol.* **2022**, *155*, 105254. [CrossRef]
13. Naveed, D.; Nadeem, H.; Sattar, M.A. Introduction of monkeypox vaccines; ahead of a looming pandemic. *Ann. Med. Surg.* **2022**. [CrossRef]

14. Khalil, A.; Samara, A.; O'Brien, P.; Morris, E.; Draycott, T.; Lees, C.; Ladhani, S. Monkeypox vaccines in pregnancy: Lessons must be learned from COVID-19. *Lancet Glob. Health* **2022**, *10*, 1230–1231. [CrossRef]
15. Harapan, H.; Wagner, A.L.; Yufika, A.; Setiawan, A.M.; Anwar, S.; Wahyuni, S.; Asrizal, F.W.; Sufri, M.R.; Putra, R.P.; Wijayanti, N.P.; et al. Acceptance and willingness to pay for a hypothetical vaccine against monkeypox viral infection among frontline physicians: A cross-sectional study in Indonesia. *Vaccine* **2020**, *38*, 6800–6806. [CrossRef]
16. Yang, Q.; Xia, D.; Syed, A.A.S.; Wang, Z.; Shi, Y. Highly accurate protein structure prediction and drug screen of monkeypox virus proteome. *J. Infect.* **2022**. [CrossRef]
17. Yin, C.; Yau, S.S.T. An improved model for whole genome phylogenetic analysis by Fourier transform. *J. Theor. Biol.* **2013**, *382*, 99–110. [CrossRef]
18. Yu, N.; Li, Z.; Yu, Z. Survey on encoding schemes for genomic data representation and feature learning—From signal processing to machine learning. *Big Data Min. Anal.* **2018**, *1*, 191–210. [CrossRef]
19. Chakravarthy, N.; Spanias, A.; Iasemids, L.D.; Tsakalis, K. Autoregressive modeling and feature analysis of DNA sequences. *EURASIP J. Adv. Signal Process.* **2004**, *1*, 952689. [CrossRef]
20. Akhtar, M.; Epps, J.; Ambikairajah, E. On DNA numerical representations for period-3 based exon prediction. In Proceedings of the IEEE International Workshop on Genomic Signal Processing and Statistics, Tuusula, Finland, 10–12 June 2007. [CrossRef]
21. Wang, S.; Yang, A. DNA solution of integer linear programming. *Appl. Math. Comput.* **2005**, *170*, 626–632. [CrossRef]
22. Cheng, W.L.; Guo, M.; Ho, M.S.H. Fast parallel molecular algorithms for DNA-based computation: Factoring integers. *IEEE Trans. NanoBioscience* **2005**, *4*, 149–163. [CrossRef]
23. Abo-Zahhad, M.; Ahmed, S.M.; Abd-Elrahman, S.A. Genomic analysis and classification of exon and intron sequences using DNA numerical mapping techniques. *Int. J. Inf. Technol. Comput. Sci.* **2012**, *8*, 22–36. [CrossRef]
24. Nair, A.S.; Sreenadhan, S.P. A coding measure scheme employing electron-ion interaction pseudopotential (EIIP). *Bioinformation* **2006**, *1*, 197–202.
25. Shen, G.; Chen, Z.; Wang, H.; Chen, H.; Wang, S. Feature fusion-based malicious code detection with dual attention mechanism and BiLSTM. *Comput. Secur.* **2022**, *119*, 102761. [CrossRef]
26. Joshi, V.M.; Ghongade, R.B.; Joshi, A.M.; Kulkarni, R.V. Deep BiLSTM neural network model for emotion detection using cross-dataset approach. *Biomed. Signal Process. Control* **2022**, *73*, 103407. [CrossRef]
27. Stallings, W.M.; Gillmore, G.M. A note on "accuracy" and "precision". *J. Educ. Meas.* **1971**, *8*, 127–129. [CrossRef]
28. Kumar, R.; Tripathi, R. A Secure and Distributed Framework for sharing COVID-19 patient Reports using Consortium Blockchain and IPFS. In Proceedings of the 6th International International Conference on Parallel, Distributed and Grid Computing, Waknaghat, India, 6–8 November 2020.
29. Bangyal, W.H.K.; Ahmad, J.; Tayyab, R.F. Optimization of Neural Network Using Improved Bat Algorithm for Data Classification. *J. Med. Imaging Health Inform.* **2019**, *9*, 670–681. [CrossRef]
30. Rukhsar, L.; Bangyal, W.H.; Khan, M.S.A.; Ibrahim, A.A.A.; Nisar, K.; Rawat, D.B. Analyzing RNA-Seq Gene Expression Data Using Deep Learning Approaches for Cancer Classification. *Appl. Sci.* **2022**, *12*, 1850. [CrossRef]