



# Article Boosted Prediction of Antihypertensive Peptides Using Deep Learning

Anum Rauf <sup>1</sup>, Aqsa Kiran <sup>2</sup>, Malik Tahir Hassan <sup>3</sup>, Sajid Mahmood <sup>2</sup>, Ghulam Mustafa <sup>2</sup>, and Moongu Jeon <sup>4,\*</sup>

- <sup>1</sup> Department of Computer Science, School of Systems and Technology (SST), University of Management and Technology (UMT), Lahore 54770, Pakistan; namraanumrauf@gmail.com
- <sup>2</sup> Department of Informatics & Systems, SST, University of Management and Technology (UMT), Lahore 54770, Pakistan; aqsa.kiran@umt.edu.pk (A.K.); sajid.mahmood@umt.edu.pk (S.M.); ghulammustafa@umt.edu.pk (G.M.)
- <sup>3</sup> Department of Software Engineering, SST, University of Management and Technology (UMT), Lahore 54770, Pakistan; tahir.hassan@umt.edu.pk
- <sup>4</sup> School of EECS, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Korea
- Correspondence: mgjeon@gist.ac.kr; Tel.: +82-62-715-2406

**Abstract:** Heart attack and other heart-related diseases are among the main causes of fatalities in the world. These diseases and some other severe problems like kidney failure and paralysis are mainly caused by hypertension. Since bioactive peptides extracted from naturally existing food substances possess antihypertensive activity, these antihypertensive peptides (AHTP) can function as prospective replacements for existing pharmacological drugs with no or fewer side effects. Such naturally existing peptides can be identified using in-silico approaches. The in-silico methods have been proven to save huge amounts of time and money in the identification of effective peptides. The proposed methodology is a deep learning-based in-silico approach for the identification of antihypertensive peptides (AHTPs). An ensemble method is proposed that combines convolutional neural network (CNN) and support vector machine (SVM) classifiers. Amino acid composition (AAC) and g-gap dipeptide composition (DPC) techniques are used for feature extraction. The proposed methodology has been evaluated on two standard antihypertensive peptide sequence datasets. The model yields 95% accuracy on the benchmarking dataset and 88.9% accuracy on the independent dataset. Comparative analysis is provided to demonstrate that the proposed method outperforms existing state-of-the-art methods on both of the benchmarking and independent datasets.

**Keywords:** convolutional neural networks (CNN); hypertension (HT); antihypertensive peptides (AHTPs); support vector machine (SVM); boosted prediction; ensemble classifier

## 1. Introduction

Since hypertension (HT) is a general medical issue that affects about twenty-five percent of the populace, the chances of a person becoming hypertensive increase with age [1]. Frequent high blood pressure generally causes hypertension and it is known as a silent killer. HT typically has no apparent symptoms, as compared to other diseases like fever and asthma. Therefore, it might take a while for a person to be diagnosed as hypertensive. Delayed diagnosis may cause severe medical issues like stroke, heart-related diseases, and other significant abnormalities like renal failure, multi-infarct dementia, harm to brain organs, and the cardiovascular illnesses [2,3].

The high prevalence and dangerous effects of hypertension signify the need to discover novel treatments and drugs to lessen/eradicate its consequences. Currently, there are many drugs for HT available on the market, like angiotensin-converting enzyme (ACE) inhibitors, beta-blockers, calcium channel blockers, k+ sparing diuretics, loop diuretics, and thiazide diuretics [4]. Although these drugs have been proven to be beneficial for



Citation: Rauf, A.; Kiran, A.; Hassan, M.T.; Mahmood, S.; Mustafa, G.; Jeon, M. Boosted Prediction of Antihypertensive Peptides Using Deep Learning. *Appl. Sci.* **2021**, *11*, 2316. https://doi.org/10.3390/ app11052316

Academic Editor: Fabrice Meriaudeau

Received: 28 January 2021 Accepted: 2 March 2021 Published: 5 March 2021

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



**Copyright:** © 2021 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/). the treatment of HT, they may cause notable side-effects, such as hypotension, metabolic alkalosis, depression, hallucinations, vivid dreams, hyperglycemia, angioneurotic edema, cough, ankle edema, tachycardia, headache, urinary urgency, functional renal insufficiency, hyponatremia, impotence, and insomnia [5]. Therefore, the discovery of secure medicines to treat and/or reduce the harmful effects of hypertension is indispensable. A medicine is considered secure if it has no or minimum side-effects.

The angiotensin-converting enzyme (ACE) is an essential element of the renin–angiotensin system (RAS). It is responsible for balancing the fluid volumes in the body and thus controlling the blood pressure. It regulates the conversion of angiotensin-I (a decapeptide) into the active angiotensin-II (an octapeptide), which constricts the blood veins. It is a powerful naphazoline hormone and a mineralocorticoid-animating peptide that controls blood pressure.

Thus, ACE is indirectly responsible for high blood pressure by causing blood veins to constrict. Many biological peptides have the potential to inhibit ACE in the renin-angiotensin system, and are thus useful to prevent and treat HT [6]. A large number of such biological peptides are found in the proteins from foods prepared from animals and plants such as, fish, cheese, milk, egg, corn, algae, microorganisms, insects, fungi, wakame, amaranth, soybean, wheat, chicken, snake, bovine, etc. [7]. Such peptides are called antihypertensive peptides (AHTPs), and their identification can lead to the development of more beneficial drugs (with less side effects) for the treatment of HT. However, the identification of a peptide that can function as an antihypertensive peptide is an expensive task in terms of time and cost. In-silico approaches can be of great help in developing systems that can absolutely filter out the non-antihypertensive peptides (non-AHTP) and provide a set of peptides with the potential to inhibit hypertensive conditions. There is a great need to develop such systems that can produce highly accurate predictions of such peptides. Recently, limited studies have demonstrated the power of machine learning (ML)-based methods to develop AHTP prediction systems.

Firstly, Wang et al. quantitatively defined the relationship among target molecular structures and biological activities, and created a quantitative structure-activity relationship (QSARs) model to confirm the structure of the ACE-inhibitor peptides and the biological activities of them. To create this model, they utilized g-scale features and used partial least square (PLS) regression-based methods. They have created this model on the basis of very small peptides (e.g., peptides with lengths of two and three), and can only predict the inhibitory activity of these tiny peptides, which is the main disadvantage of this study [8]. Another approach was built by Kumar et al. in 2015 in this area [9]. The dataset was divided into four different categories to extract features. The categories were named (1) tiny (dipeptides and tripeptides), (2) small (tetrapeptides, pentapeptides, and hexapeptides), (3) medium (i.e., sizes ranging between seven and twelve), and (4) large peptides (greater than twelve amino acids). For tiny peptides, chemical descriptors were extracted to generate support vector machine (SVM)-based regression models, and they achieved a correlation of 0.701 for dipeptides and 0.543 for tripeptides. For smaller peptides, SVM-based classification models were developed, and the maximum obtained accuracies were 76.67% for tetrapeptide, 72.04% for pentapeptide, and 77.39% for hexapeptides. For medium and large peptides, amino acid composition features were extracted to develop SVM-based classification models, and accuracies of 82.61% for medium peptides and 84.21% for large peptides were attained. Moreover, a web-based platform called AHTpin was also created to screen, predict, and design AHTPs. Win et al. developed a computerized AHTPs prediction system [10] that employs random forest (RF) models. The models are trained by using the groupings of amino acid composition, dipeptide composition, and pseudo amino acid composition feature encoding techniques. The system demonstrates a marginal improvement over AHTpin, with an accuracy of 84.73% when run on an independent test dataset. Furthermore, the feature importance analysis emphasized the preference at the C-terminal of the proline amino acids and non-polar amino acids, and also the capacity of little peptides for vigorous activity.

An online web platform (called PAAP) is also available for public use of the proposed model it contains. The mAHTpred meta-predictor is another approach to classifying AHTPs [11]. To identify AHTPs, Manavalan et al. developed mAHTpred by using eight feature encoding schemes to construct 51 feature vectors from two different datasets (benchmarking dataset and independent dataset). Extremely randomized tree (ERT)-based models were created by using these 51 feature vectors. A new feature vector consisting of the predicted probabilities of AHTPs was calculated by using the above mentioned ERT-based models. The new feature vector was then utilized as an input for four different ML algorithms: SVM, gradient boosting (GB), RF, and ERT. Final predictions were made by using an ensemble of SVM, GB, RF, and ERT models.

AHTP prediction has been realized in [12], using recursive feature elimination. This work generates and feeds optimal features into the ensemble of four classification algorithms (SVM, C4.5 Decision Tree, random forest (RF), and extreme gradient boosting (XGBoost)) in order to achieve the final prediction. Deep-AmPEP30 is a recent deep learning-based model, devised for the prediction of short-length antimicrobial peptides (AMP), which is another important bioactive peptide sequence [13]. Deep-AmPEP30 employs CNN on a reduced set of amino acid composition (AAC) features. This group experimented with a benchmark dataset consisting of balanced classes, with accuracy scores of 77% and 85% gained by both Area under the Receiver Operating Characteristic Curve (AUC-ROC) and Area under the Precision-Recall Curve (AUC-PR), respectively. In this research activity, a deep learning-based antihypertensive peptide predictor is presented. The proposed approach uses the standard datasets used in [9] and [11]. Features are generated by using two feature-encoding techniques: amino acid composition (AAC) and g-gap dipeptide composition (g-gap DPC). Dipeptide composition features are further represented as red, green, and blue (RGB) images. An RGB image is created against each dataset. In the literature, ensemble-based methods have been successfully applied for the task of classification [11,12]. In such methods, multiple classifiers are combined in parallel or in sequence to boost the classification performance. In the proposed method as well, an ensemble of classifiers is used to make a boosted prediction by employing a convolutional neural network (CNN) followed by a support vector machine (SVM). Four different CNN models are trained on the generated image dataset. The predicted outputs of these four CNN models are combined with the AAC features for every sequence, resulting in new feature vectors. These new feature vectors are then used to train an SVM model for the final classification of the peptide sequences as either antihypertensive or not. The proposed predictor is evaluated using a 10-fold cross-validation method, and achieves an accuracy of 95% on the benchmarking dataset and 88.9% on the independent dataset.

# 2. Materials and Methods

This work is carried out in the following six steps: (1) data acquisition and analysis, (2) feature extraction, (3) RGB image generation, (4) training the CNN models, (5) generating new feature vectors on the basis of the CNN model's output and the ACC feature, and (6) training the SVM classifier. Details of each step are provided in this section. The overall methodology is outlined in Figure 1.

# 2.1. Data Acquisition and Analysis

Two datasets are used in this work. One is considered as the benchmarking dataset [9] and the other one is taken as an independent dataset [11]. We used the same benchmarking dataset during the model construction as was used in previous works [9,11] to compare our approach. The independent dataset was used for test data during the evaluation. Both datasets consist of antihypertensive peptide (AHTP) sequences as positive class samples and non-antihypertensive peptide (non-AHTP) sequences as negative class samples. All peptide sequences are made up of twenty amino acid residues. The benchmarking dataset contains 913 samples of the positive class and 913 samples of the negative class. Similarly, the independent dataset consists of 386 samples of each class. Data analysis of both datasets,

and the total number of occurrences of each of the twenty amino acid residues (A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, V) in both datasets, is given in Table 1.



Figure 1. Steps of the proposed methodology.

Statistics for the peptide sequences are presented in Table 2. It shows the minimum, maximum and average lengths of the peptide sequences in both datasets.

## 2.2. Feature Extraction

Features are extracted from the peptide sequence dataset by applying two feature encoding schemes (i.e., AAC and g-gap DPC). In total, 20 features are extracted by using AAC and 400 features (of shape 20-by-20) by using each DPC with a gap of zero, one, two, and three, respectively. Overall, there are 1620 features for each sample of the dataset. The feature encoding schemes are described in detail below.

## 2.2.1. Amino Acid Composition (AAC)

There are twenty standard amino acids that repeat in each protein sequence. Amino acid composition is a feature extraction technique that represents the peptide by calculating the percentage of each amino acid in a given peptide sequence. Figure 2 shows an example of calculating the AAC features for a sample peptide sequence. In this way, we get a feature vector of size twenty.

# 2.2.2. G-Gap Dipeptide Composition (DPC)

A dipeptide is a composite of two amino acid residues, with or without a gap between them. DPC is computed as the ratio of the number of occurrences of a dipeptide to the entire length of the sequence. It can also be computed with a gap of amino acids, where the value of g ranges from 0 to 3. Such a DPC feature is called g-gap dipeptide composition. DPC generates a 20-by-20 matrix that contains 400 features for each value of the g-gap, as shown in Figure 3 (for g-gap = 0).

Dataset	Α	R	Ν	D	С	Q	Ε	G	Н	Ι	L	К	Μ	F	Р	S	Т	W	Y	V
Independent Positive	164	99	46	62	20	120	103	157	82	133	225	164	99	46	62	20	120	103	157	82
Independent Negative	319	270	337	306	280	276	297	304	247	301	341	319	270	337	306	280	276	297	304	247
All Independent	483	369	383	368	300	396	400	461	329	434	566	483	369	383	368	300	396	400	461	329
Benchmarking Positive	374	263	236	174	48	410	354	442	202	368	625	374	263	236	174	48	410	354	442	202
Benchmarking Negative	607	420	331	404	77	305	454	547	149	414	687	607	420	331	404	77	305	454	547	149
All Benchmarking	981	683	567	578	125	715	808	989	351	782	1312	981	683	567	578	125	715	808	989	351

Table 1. Occurrence of each amino acid residue in both datasets.



Figure 2. Example of amino acid composition (AAC) features for sample peptide sequence.

|--|

Datasets	Min Length of Sequence	Max Length of Sequence	Avg. Length of Sequence
Benchmarking   (Negative Class)	5	45	8.048
Benchmarking (Positive Class)	5	81	7.746
ALL Benchmarking	5	81	7.897
Independent (Negative Class)	5.0	29.0	6.48
Independent (Positive Class)	5.0	24.0	15.42
All Independent	5.0	29.0	10.95



Figure 3. Dipeptide composition feature of a sample peptide with a gap of 0.

# 2.3. RGB Image Generation

DPC-generated 20-by-20 matrices are used to generate RGB images. Each image consists of three ( $20 \times 20$ ) matrices. The values of each matrix are taken as pixels and the three matrices represent the three colors: red, green, and blue (RGB). Matrices with different combinations of g-gaps are used to generate distinct images for each g-gap, as given in Table 3. For example, the first image is generated by considering a DPC with g-gap = 0 as the red color matrix, a DPC with g-gap = 1 as the green color matrix, and a DPC with g-gap = 2 as the blue color matrix.

Table 3. Combinations of g-gap dipeptide composition (DPC) features for the sample peptide.

Model	Combinations
G013	DPC with gap 0, DPC with gap 1, DPC with gap 3
G012	DPC with gap 0, DPC with gap 1, DPC with gap 2
G023	DPC with gap 0, DPC with gap 2, DPC with gap 3
G123	DPC with gap 1, DPC with gap 2, DPC with gap 3

Figure 3 shows that the values in all matrices are less than 1, and thus could not be properly visualized in the image. To visualize the image properly, we have replaced all the non-zero values of DPC features with 255 (i.e., the largest value of pixel). Figure 4 shows the procedure for converting the peptide to an image for the first combination on the sample peptide. Using this procedure, we generated four image datasets, as shown in Figure 5.



Figure 4. Conversion of sample peptide to image for first combination.



Figure 5. Four types of RGB images generated against the sample protein sequence.

#### 2.4. Trainig CNN Models

In this step, the convolutional neural network (CNN) is used, which is one of the most prominently utilized approaches for image analysis [14]. It first extracts features from an image by applying convolutional layers, i.e., by striking the number of feature maps (kernels) on it. These feature maps are considered as features of the image, and then pooling layers are applied to reduce the dimensionality of each feature map without losing useful information. This is followed by a classification (dense) layer, which is a conventional artificial neural network [14].

In the proposed approach, each image generated in the previous step 2.3 is used as an input for the CNN. Four CNN models are generated for the classification of each sample peptide. The CNN architecture and parameters used are defined in Table 4 and SVM parameters are defined in Table 5. We take the prediction value of each four CNN models, which is either 0 if the predicted class is negative or 1 if the predicted class is positive. The performance evaluation of CNN models is presented in Table 6. Although the CNN models demonstrate reasonable accuracy, the performance was inferior to the benchmark approaches in [9] and [11]. Thus, we apply a boosted prediction approach by combining another promising classifier (SVM) which outperforms the benchmark approaches. SVM has gained prominent success in bioinformatics, including in the classification of protein sequences [15,16]. A proposed boosting step is described in the following section.

Table 4. Parameters for CNN.

Parameters	Values
Input shape	20  imes 20  imes 3
Number of convolutional layers	2
Number of dense layers	2
Number of filters in first convolutional layer	32
Number of filters in second convolutional layer	8
Shape of filters	$3 \times 3$
Maxpooling shape	2  imes 2
Dropout rate	0.5
Batch size	35
Epochs	270
Activation functions in convolutional layers	Relu, Sigmoid
Optimizer	Adam
Loss function	Binary Crossentropy
Number of filters in first dense layer	170
Number of filters in second dense layer	1
Activation functions in dense layers	Relu, Sigmoid

Parameters	Values
Kernel	Radial Basis Function
Gemma	0.5
Constant (c)	2

Table 5. Parameters used in the support vector machine (SVM) classifier.

**Table 6.** The 10-fold cross-validation results of 4 CNN models on benchmarking and independent datasets.

Model	Dataset	ACC	Sp	Sn	MCC	AUC
G013		0.8226	0.7974	0.8477	0.6477	0.8226
G012	Pon chan arlin a	0.8099	0.7985	0.8213	0.6217	0.8099
G023	benchinarking	0.8078	0.7853	0.8302	0.6168	0.8078
G123		0.8083	0.7865	0.8306	0.6185	0.8083
G013		0.8696	0.9221	0.9454	0.8696	0.8696
G012	Independent	0.9352	0.9326	0.9377	0.8708	0.9352
G023		0.9350	0.9298	0.9402	0.8706	0.9350
G123		0.9324	0.9195	0.9453	0.8665	0.9324

2.5. Boosting as Input Prediction by Ensembling an SVM Classifier Using a Combination of CNN Models as Output and AAC Features as Input

A new feature vector is generated by combining the CNN models' outputs and the 20 AAC features. In this way, we get a new feature vector against each sample. The length of this feature vector is 24.

These 24 features are presented to the SVM for final classification. The parameters used for the SVM classifier are presented in Table 5. The performance evaluation is discussed in Section 3.1 and Table 7. The obtained results show that the proposed method using boosting outperforms the CNN models as well as the existing approaches.

**Table 7.** The 10-fold cross-validation results of the boosted-SVM predictor on the benchmarking and independent datasets.

Dataset	ACC	Sp	Sn	MCC	AUC
Benchmarking	0.9584	0.9201	0.9967	0.9203	0.9584
Independent	0.9235	0.8471	1.0	0.8576	0.9235

### 2.6. Computational Environments

The proposed approach was developed on Google Colaboratory (which is a cloudbased online Jupyter Notebook platform) by using python programming language. We chose Google Colaboratory because we did not have to install python and its libraries manually, and it has a diverse range of python libraries already installed on the cloud. It also provides Graphics Processing Unit (GPU), which makes the training process relatively fast. For deep learning, we have used python's library Keras with the TensorFlow backend.

#### 3. Results and Discussion

The experimental results of the proposed approach are discussed in this section. A comparative analysis with existing methods is also provided to demonstrate that the proposed approach outperforms the existing approaches.

#### 3.1. Model Evaluation Results on Benchmarking and Independent Dataset

The proposed model is evaluated on both benchmarking and independent datasets, by using the 10-fold cross validation. Commonly used performance metrics are calculated. The metrics include accuracy (ACC), sensitivity (Sn), specificity (Sp), Matthews's correlation

coefficient (MCC), and area under the curve (AUC). Four CNN models are trained on both datasets, and the results of the performance metrics are shown in Table 6. The outputs of the CNN models and the AAC features are then used as inputs to train the SVM model, which is our final predictor. Table 7 shows the performance of the SVM model on both datasets.

# 3.2. Comparative Analysis and Discussion

The performance of this work is compared with three existing approaches, including AHTpin [9,10] and mAHTPred [11]. A detailed description of the existing approaches is provided in the introduction section. AHTpin consists of two prediction models; one is based on AAC features and the other is based on atomic composition features. Our results are compared with both of these two models. mAHTPred achieved the highest accuracy as compared to the previous two techniques [11]. We compared our results with those of both mAHTPred and the other two techniques, by running them on the same datasets.

The comparison results show that the proposed boosted predictor outperforms the existing techniques, as shown in Table 8. The boosted predictor for AHTPs achieved a better performance on both datasets, in terms of ACC, Sp, Sn, MCC, and AUC. The ACC and MCC results are approximately 9–21% higher than those of the previous approaches.

**Table 8.** Comparison of the proposed approach and existing approaches on benchmarking and independent datasets.

Methods	Dataset	ACC	Sp	Sn	MCC	AUC
Proposed		0.958	0.920	0.996	0.920	0.958
mAHTPred		0.848	0.874	0.821	0.697	0.865
PAAP	Benchmarking	0.791	0.780	0.865	0.585	NA
AHTpin_AAC	0	0.785	0.793	0.865	0.585	NA
AHTpin_ATC		0.785	0.787	0.865	0.573	NA
Proposed		0.895	0.841	0.948	0.795	0.895
mAHTPred		0.883	0.873	0.894	0.767	0.951
PAAP	Independent	NA	NA	NA	NA	NA
AHTpin_AAC	-	0.800	0.800	0.800	0.601	0.800
AHTpin_ATC		0.798	0.842	0.798	0.641	0.888

The independent dataset was created by Manavalan et al. and we used it to check the robustness of mAHTPred [11]. Comparison results of independent dataset are also provided in Table 8.

# 4. Conclusions

Hypertension is connected to numerous diseases such as cancer, heart attack, renal failure, and paralysis. Bioactive peptides are derived naturally and have antihypertensive activity. Bioactive peptides work as encouraging substitutes to pharmacological medicines. Such peptides are useful, but to find out whether the peptides have antihypertensive characteristics or not is an expensive process. Bioinformatics could be used to build an automated system to identify the antihypertensive peptides and some solutions have already been proposed [9,11]. However, there is still room for performance improvement in such tools. In this paper, we present an automated antihypertensive peptides prediction system. It takes a peptide sequence as the input and predicts whether the given peptide is antihypertensive or not. The comparative results demonstrate that the proposed approach outperforms the existing state-of-the-art approaches. Experimental results show that the proposed approach yields high accuracies on standard datasets, as compared to the previous approaches.

In future, the proposed approach can be applied to other types of biological datasets. Exploration of other deep learning techniques, performance optimization, and the launching of a web service are also promising future directions. **Author Contributions:** The idea conceptualization, data curation and methodology design were done by M.T.H. and S.M. Original draft was prepared by A.R. and this was proofread by A.K. Implementation was performed by A.R. and validation was done by A.K. Project administration was done by G.M. The review and editing were performed by G.M. and M.J. Funding was acquired by M.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by GIST Research Project grant funded by the GIST in 2021. Support of UMT is also acknowledged.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

**Data Availability Statement:** The datasets used in this study are public datasets already used by other researchers e.g., [9] and [11].

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

# References

- 1. Chockalingam, A.; Campbell, N.R.; Fodor, J.G. Worldwide epidemic of hypertension. Can. J. Cardiol. 2006, 22, 553–555. [CrossRef]
- Cunha, J.P.W.; Marks, J. High Blood Pressure (Hypertension). 2011. Available online: http://www.medicinenet.com/high\_ blood\_pressure/article.htm (accessed on 24 November 2020).
- Fisher, N.D.L.; Curfman, G. Hypertension—A Public Health Challenge of Global Proportions. JAMA 2018, 320, 1757–1759. [CrossRef] [PubMed]
- 4. Zisaki, A.; Miskovic, L.; Hatzimanikatis, V. Antihypertensive Drugs Metabolism: An Update to Pharmacokinetic Profiles and Computational Approaches. *Curr. Pharm. Des.* 2014, *21*, 806–822. [CrossRef] [PubMed]
- 5. Laurent, S. Antihypertensive drugs. *Pharmacol. Res.* 2017, 124, 116–125. [CrossRef] [PubMed]
- Hong, F.; Ming, L.; Yi, S.; Zhanxia, L.; Yongquan, W.; Chi, L. The antihypertensive effect of peptides: A novel alternative to drugs? *Peptides* 2008, 29, 1062–1071. [CrossRef] [PubMed]
- Kumar, R.; Chaudhary, K.; Sharma, M.; Nagpal, G.; Chauhan, J.S.; Singh, S.; Gautam, A.; Raghava, G.P. AHTPDB: A comprehensive platform for analysis and presentation of antihypertensive peptides. *Nucleic Acids Res.* 2014, 43, D956–D962. [CrossRef] [PubMed]
- 8. Wang, X.; Wang, J.; Lin, Y.; Ding, Y.; Wang, Y.; Cheng, X.; Lin, Z. QSAR study on angiotensin-converting enzyme inhibitor oligopeptides based on a novel set of sequence information descriptors. *J. Mol. Model.* **2010**, *17*, 1599–1606. [CrossRef] [PubMed]
- 9. Kumar, R.; Chaudhary, K.; Chauhan, J.S.; Nagpal, G.; Kumar, R.; Sharma, M.; Raghava, G.P. An in silico platform for predicting, screening and designing of antihypertensive peptides. *Sci. Rep.* **2015**, *5*, srep12512. [CrossRef] [PubMed]
- 10. Win, T.S.; Schaduangrat, N.; Prachayasittikul, V.; Nantasenamat, C.; Shoombuatong, W. PAAP: A web server for predicting antihypertensive activity of peptides. *Future Med. Chem.* **2018**, *10*, 1749–1767. [CrossRef] [PubMed]
- 11. Manavalan, B.; Basith, S.; Shin, T.H.; Wei, L.; Lee, G. mAHTPred: A sequence-based meta-predictor for improving the prediction of anti-hypertensive peptides using effective feature representation. *Bioinformatics* **2019**, *35*, 2757–2765. [CrossRef] [PubMed]
- 12. Chang, W.; Liu, Y.; Xiao, Y.; Yuan, X.; Xu, X.; Zhang, S.; Zhou, S. A machine learning-based Prediction method for hypertension outcomes based on medical data. *Diagnostics* **2019**, *9*, 178. [CrossRef] [PubMed]
- Yan, J.; Bhadra, P.; Li, A.; Sethiya, P.; Qin, L.; Tai, H.K.; Wong, K.H.; Siu, S.W. Deep-AmPEP30: Improve Short Antimicrobial Peptides Prediction with Deep Learning. *Mol. Ther. Nucleic Acids* 2020, 20, 882–894. [CrossRef] [PubMed]
- 14. Stenroos, O. Object Detection from Images Using Convolutional Neural Networks. Master's Thesis, 2017. Available online: https://aaltodoc.aalto.fi/bitstream/handle/123456789/27960/master\_Stenroos\_Olavi\_2017.pdf?sequence=1&isAllowed=y. (accessed on 11 November 2020).
- Ho Thanh Lam, L.; Le, N.H.; Van Tuan, L.; Tran Ban, H.; Nguyen Khanh Hung, T.; Nguyen, N.T.K.; Dang, L.H.; Le, N.Q.K. Machine learning model for identifying antioxidant proteins using features calculated from primary sequences. *Biology* 2020, 9, 325. [CrossRef] [PubMed]
- 16. Manavalan, T.B.; Shin, H.; Lee, G. PVP-SVM: Sequence-based prediction of for improving the prediction of Phage Virion proteins using a Support Vector Machine. *Front. Microbiol.* **2018**, *9*, 476. [CrossRef]