



Klasifikasi Coronary Heart Disease (CHD) Berbasis Optimasi DNN dan Inisialisasi Kaiming He

Lia Andiani*, Sukemi, Dian Palupi, Nurul Afifah

Fakultas Ilmu Komputer, Universitas Sriwijaya, Palembang, Indonesia

Email: ^{1,*}liaandiani199@gmail.com, ²sukemiku66@gmail.com, ³dianpalupirini@gmail.com, ⁴nurulafifah@gmail.com

Email Penulis Korespondensi: liaandiani199@gmail.com

Abstrak—CHD adalah nyeri dada atau ketidaknyamanan yang terjadi jika area otot jantung tidak mendapatkan cukup darah yang kaya oksigen. CHD juga dikenal sebagai penyakit arteri koroner. CHD meningkat setiap tahun dengan jumlah kematian yang signifikan. Diusulkan algoritma pembelajaran untuk mendapatkan kinerja yang lebih baik dalam akurasi, sensitivitas, dan spesifisitas dalam interpretasi CHD. Akurasi dapat ditingkatkan dengan menambahkan suatu teknik optimasi inisialisasi weight Kaiming He (2015) pada struktur DNN. Maka dari itu kami mengusulkan DNN di optimasi dengan suatu teknik inisialisasi weight Kaiming He sehingga dapat mengatasi kelemahan dalam varian data. Ini dibuktikan dengan hasil performa akurasi sebesar 98.73%. Inisialisasi bobot kaiming he terbukti dapat meningkatkan akurasi dan mengatasi masalah varian data yang besar antar kelas.

Kata Kunci: CHD; Kaiming He; DNN; Akurasi

Abstract—CHD is chest pain or discomfort that occurs if the area of the heart muscle does not get enough oxygen-rich blood. CHD is also known as coronary artery disease. CHD is increasing every year with a significant number of deaths. A learning algorithm is proposed to get better performance in accuracy, sensitivity, and specificity in CHD interpretation. Accuracy can be improved by adding a Kaiming He (2015) weight initialization optimization technique to the DNN structure. Therefore we propose that DNN is optimized with a Kaiming He weight initialization technique so that it can overcome weaknesses in the data variant. This is evidenced by the results of the accuracy performance of 98.73%. Initialization of kaiming he weights is proven to improve accuracy and overcome the problem of large data variants between classes.

Keywords: CHD; Kaiming He; DNN; Accuracy; Sensitivity; Specificity

1. PENDAHULUAN

Penyakit pada organ jantung atau kardiovaskular adalah penyebab nomor satu kematian pada 17,7 juta kematian dalam penelitian[1]. Di Indonesia, presentasi tertinggi penyakit kardiovaskular adalah penyakit serangan jantung koroner (*Coronary Heart Disease*) sebesar 1,5 persen. CHD yang diam adalah orang yang menderita tetapi tidak memiliki gejala. Penyakit ini mungkin tidak didiagnosis sampai seseorang memiliki gejala gagal jantung atau serangan jantung[2], menggunakan teknik pemodelan *data mining*, yaitu, *Naïve Bayes* untuk memprediksi penyakit jantung dengan menghasilkan akurasi sebesar 82,31% [2]. menyajikan sejumlah teknik kecerdasan komputasi dalam pendeteksian penyakit jantung menggunakan UCI *Machine Learning Repository* dengan menghasilkan akurasi sebesar 86.77% [3]. Teknik *data mining* untuk menyelidiki penyakit jantung menggunakan 11 atribut dari UCI *Machine Learning*. Pada publikasi Shinde dkk[4]. mengusulkan *neural network feedforward multilayer* dan *neural network backpropagation* dengan tiga lapisan tersembunyi yang terdiri dari total 18 node untuk mendiagnosis penyakit jantung. Keakuratan yang dicapai melalui sistem yang diusulkan adalah 92%. Dataset penyakit jantung *Cleveland* memberikan beberapa informasi tentang tipe nyeri dada (*angina*) yang menjadi gejala CHD yang signifikan. Dalam pekerjaan sebelumnya, beberapa algoritma pembelajaran mesin digunakan dalam interpretasi CHD[5]. Samanta dkk mengusulkan jaringan saraf tiruan (*Neural Network*) dan akurasi keseluruhan yang diperoleh adalah 75%, dengan satu lapisan tersembunyi[5]. Ramesh dkk mengusulkan *Naïve Bayes* dan peningkatan akurasi yang dicapai adalah 82,31% [6]. Nahar dkk mengusulkan teknik kecerdasan komputasi untuk prediksi penyakit jantung dan mendapatkan akurasi 86,77% [7]. Adnan dkk mengusulkan *neural network* dengan satu lapisan tersembunyi dan menghasilkan akurasi 85%. Pada publikasi [1] jumlah populasi CHD meningkat setiap tahun dengan jumlah kematian yang signifikan, Diusulkan algoritma pembelajaran untuk mendapatkan kinerja yang lebih baik dalam akurasi, sensitivitas, dan spesifisitas dalam interpretasi CHD. Sistem klasifikasi berbasis fitur tradisional kurang efektif karena kinerjanya biasanya tergantung pada kualitas fitur [1] Xing Z dkk sehingga *neural network* bisa menjadi solusi dalam penyelesaian masalah ini. [8] Pada publikasi Dharmawahyuni dkk dan Kathlen & Julia [8] jumlah varian pada kelas dataset sangat besar sehingga menjadi masalah dalam hal akurasi. Akurasi dapat ditingkatkan dengan menambahkan suatu teknik optimasi inisialisasi *weight Kaiming He* [9] pada struktur DNN.

Penelitian dalam beberapa tahun terakhir yang membahas CHD seperti dari penelitian Sundar dkk [10] yang membahas analisis performa pada data mining dalam kasus CHD[4]. Penelitian Shinde dkk membahas CHD dalam prediksi menggunakan sistem *multi layer perceptron* pada sistem *backpropagasi neural network*. Penelitian Deng dkk [11] membahas tentang cara meningkatkan performa *confusion matrix* dalam sistem klasifikasi menggunakan *neural network*. Penelitian Dahl dkk [12] membahas cara meningkatkan performa dalam sistem klasifikasi yang menggunakan algoritma DNN.

Penelitian Dharmawahyuni dkk [1] membahas sistem klasifikasi CHD menggunakan DNN. Penelitian Kathlen & Julia [8] membahas bagaimana mengklasifikasi CHD menggunakan DNN namun, akurasi dalam



publikasi ini masih kecil sehingga diperlukan suatu optimasi untuk meningkatkan akurasi. Penelitian Patel dkk [13] membahas cara menambahkan teknik inisialisasi weight dalam neural network untuk meningkatkan performa sistem klasifikasi. Penelitian Bhatia dkk [14] membahas bagaimana efek yang diberikan jika suatu sistem klasifikasi *neural network* diberikan optimasi *inisialisasi weight* menggunakan fungsi aktivasi sigmoid. Penelitian Siddharth dkk [13] membahas bagaimana cara menerapkan inisialisasi weight dalam sistem klasifikasi yang menggunakan algoritma DNN. Penelitian Kaiming He[9] membahas bagaimana cara implementasi inisialisasi weight Kaiming He dalam sistem klasifikasi yang menggunakan *neural network*. Maka dari itu diusulkan DNN di optimasi dengan suatu teknik inisialisasi weight Kaiming He sehingga dapat mengatasi kelemahan dalam varian data [15].

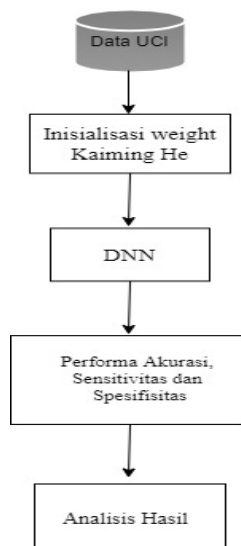
2. METODOLOGI PENELITIAN

Pada proses ini dilakukan persiapan data. Dataset yang digunakan merupakan dataset UCI serangan jantung coroner. Data CHD Present yaitu orang yang terkena serangan jantung berjumlah 815 dan CHD Absent yaitu orang sehat yang tidak terkena serangan jantung berjumlah 211 yang bisa dilihat pada tabel 2.1. Pada tahapan ini dilakukan proses pengolahan data untuk mempersiapkan data sebelum dimasukkan kedalam proses klasifikasi. Pada penelitian ini menggunakan klasifikasi biner untuk CHD *present* dan CHD *absent*. Jumlah data yang digunakan berjumlah 1026 data. Atribut dataset jantung tersaji pada Tabel I.

Tabel 1. Atribut Dataset Jantung

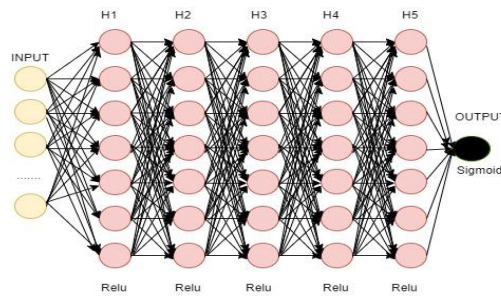
Age	Usia
Sex	Jenis Kelamin (1 = male; 0 = female)
Cp	Tipe nyeri dada
Trestbps	Tekanan darah
Chol	Kolesterol
Fbs	Gula Darah
Restcgs	Hasil Elektrokardiogram
Thalach	Denyut Jantung
Exang	Induksi Angina
Oldpeak	Depresi
Slope	Kemiringan ST
Ca	Jumlah vessel
Thai	Diagnosis
Target	0 = CHD absent, 1 = CHD present

Alur kerja penelitian yang dilakukan yaitu dengan mengumpulkan data dari UCI Repository, lalu menginisialisasi weight Kaiming He dan diteruskan dengan klasifikasi menggunakan DNN sehingga didapat performa hasil yang diinginkan. Terakhir lalu proses analisis. Berikut adalah research methodology seperti pada gambar 1.



Gambar 1. Research Methodology

Setelah melakukan research methodology, hal yang dilakukan selanjutnya yaitu merancang arsitektur DNN yang akan digunakan. Berikut arsitektur DNN yang digunakan seperti pada gambar 2.



Gambar 2. Arsitektur DNN

Berdasarkan Gambar 2 arsitektur DNN yang dibuat memiliki *input*, *layer hidden* dan *output*. *input* terdiri dari 14 atribut. Ada 5 *layer hidden* dengan menggunakan fungsi aktivasi ReLu. Untuk *hidden* pertama terdiri dari 200 *neuron*. *Layer hidden* kedua terdiri dari 100 *neuron*. *hidden* ketiga 80 *neuron*. *hidden* keempat 60 *neuron*. *hidden* kelima 20 *neuron* dan *layer output* satu *neuron* dikarenakan fungsi aktivasi pada *output* adalah sigmoid dan algoritma yang digunakan adalah *backpropagation*. Saat proses pelatihan data, jumlah data akan dibagi yaitu jumlah data untuk proses pelatihan (*training*) adalah 70% dan 30% untuk pengujian (*testing*).

Berikut algoritma dalam proses klasifikasi CHD *present* dan CHD *absent* menggunakan DNN:

- Input* data dari dataset orang sakit dan sehat berdasarkan atribut yang ada
- Inisialisasi weight dengan menggunakan persamaan *Kaiming He*. Inisialisasi weight *Kaiming He* yaitu dengan cara menginisialisasi menggunakan persamaan *Kaiming He* pada bab 2. Hasil dari perhitungan dari persamaan tersebut akan menjadi standar deviasi yang akan menjadi batasan dalam inisialisasi nilai weight. Hasil dari inisialisasi weight akan masuk ke proses *feedforward* dan *backward*. Nilai weight tersebut akan stabil karena dalam batasan *range* nilai. Hal itu yang dapat mengatasi masalah varian data yang besar.

$$\sigma = \sqrt{\frac{2}{n_{in}}} \tag{1}$$

- Proses *Learning data* dimulai dari proses *feedforward* yaitu menerima input x_i dan diteruskan ke *layer* selanjutnya yaitu *hidden layer*. *Hidden layer* akan menghitung fungsi aktivasinya dan mengirimkan sinyal ke *output*. Fungsi aktivasi yang digunakan adalah relu dan sigmoid, maka dari *hidden layer* 1 hingga 5 menggunakan fungsi aktivasi relu dan pada *output* menggunakan sigmoid.
- Loss* digunakan untuk mengupdate weight pada proses *backward* melalui proses *chain rule*. Berikut turunan parsial *output* yang akan terus diturunkan hingga *input layer* untuk membuat *loss* semakin kecil.

$$\frac{\partial E}{\partial y_i} = -1 \times \left(out \times \frac{1}{y_i} \right) + (1 - out) \times \left(\frac{1}{1-y_i} \right) \tag{2}$$

- Update weight dan bias

$$W_{baru} = W_{lama} + \Delta W \tag{3}$$

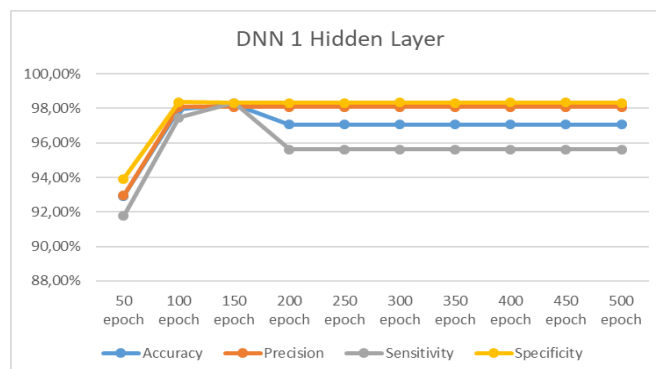
$$b_{baru} = b_{lama} + \Delta W \tag{4}$$

- Simpan weight

Klasifikasi data sehingga dihasilkan dua kelas label yaitu CHD *present* dan CHD *absent*.

3. HASIL DAN PEMBAHASAN

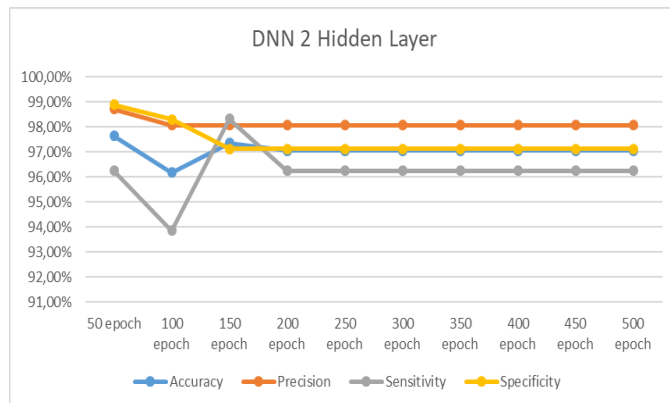
Hasil uji validasi dari DNN dengan 1 hidden layer yang dioptimasi dengan Kaiming He seperti pada gambar 3.



Gambar 3. DNN 1 Layer

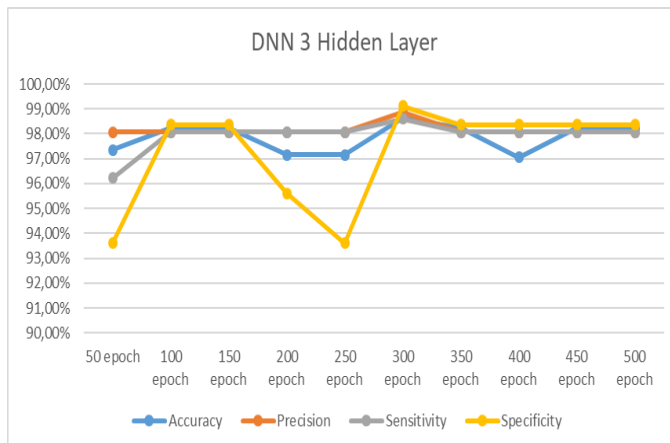


Pada gambar 3 merupakan hasil validasi DNN 1 layer menunjukkan pada percobaan 150 epoch menghasilkan akurasi terbaik yaitu 98.23% dari seluruh pengujian DNN 1 layer. Pada pengujian 150 epoch, bobot yang diinisialisasi cukup baik sehingga bias membuat nilai akurasi menjadi tinggi. Seperti pada gambar 4.



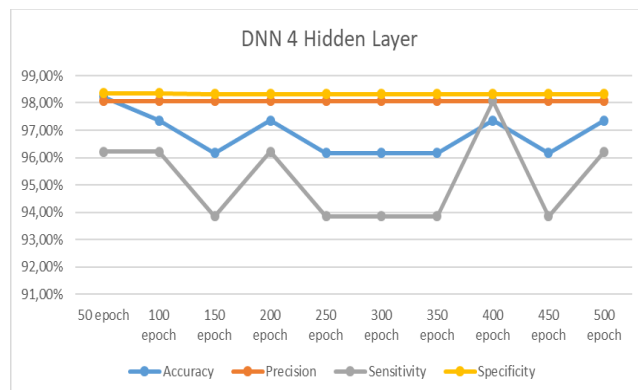
Gambar 4. DNN 2 Layer

Pada gambar 4 hasil validasi DNN 2 layer menunjukkan pada percobaan 50 epoch menghasilkan akurasi terbaik yaitu 97.64% dari seluruh pengujian DNN 2 layer. Pada pengujian 50 epoch, bobot yang diinisialisasi tidak cukup baik dibandingkan inisialisasi bobot DNN dengan 1 hidden layer dikarenakan akurasi tertinggi dari validasi pengujian ini tidak lebih baik dari pada DNN dengan 1 hidden layer. seperti pada gambar 5.



Gambar 5. DNN 3 Layer

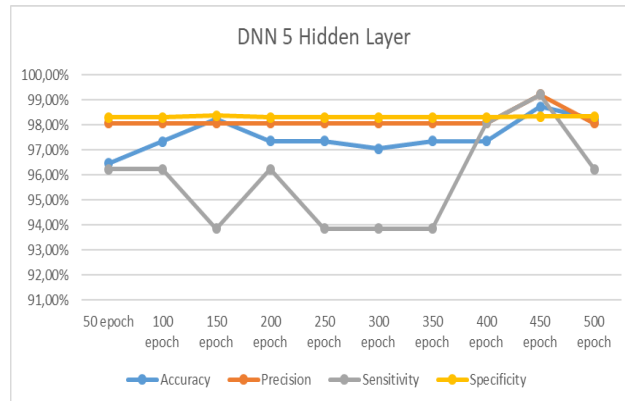
Pengujian DNN dengan 3 hidden layer menghasilkan akurasi tertinggi pada 300 epoch yaitu dengan persentase akurasi sebesar 98.65%. Presisi tertinggi yaitu 98.87%. Sensitivitas sebesar 98.60% dan spesifisitas 99.12%. DNN dengan 3 hidden layer menghasilkan nilai terbaik. Bobot yang diinisialisasi menghasilkan nilai bobot dengan rentang nilai berdekatan sehingga varians data dapat diatasi dibuktikan dengan performa akurasi, presisi, sensitivitas dan spesifisitas yang tinggi. Seperti pada gambar 6.



Gambar 6. DNN 4 Layer

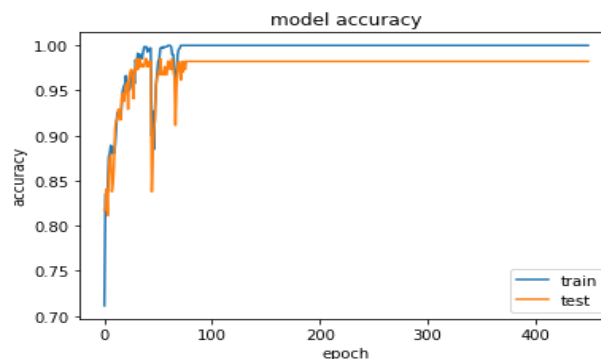


Pengujian validasi DNN dengan 4 hidden layer, akurasi tertinggi diperoleh saat 50 epoch dengan persentase akurasi sebesar 98.23%. namun hasil akurasi yang dihasilkan tidak terlalu baik dari hasil validasi DNN dengan 3 hidden layer. Bobot yang diinisialisasi menghasilkan angka dengan rentang yang cukup jauh sehingga model DNN dengan 4 hidden layer tidak terlalu baik dalam hal performa akurasi, sensitivitas, presisi dan spesifisitas. Seperti pada gambar 7.



Gambar 7. DNN 5 Layer

Pengujian DNN dengan 5 hidden layer menghasilkan akurasi tertinggi saat validasi 450 epoch. Presisi dengan persentase 99.21%, sensitivitas sebesar 99.11% dan spesifisitas sebesar 98.36%. bobot yang diinisialisasi memiliki rentang yang baik sehingga hasil performa mendapat hasil yang bagus. Seluruh validasi yang sudah diuji dapat dianalisis bahwa DNN dengan 5 hidden layer menjadi model terbaik dalam pengujian. Ini dibuktikan dengan hasil performa akurasi sebesar 98.73%. Inisialisasi bobot kaiming he terbukti dapat meningkatkan akurasi dan mengatasi masalah varian data yang besar antar kelas. Hasil Validasi dari DNN dengan menggunakan 5 hidden layer selama 450 epoch menghasilkan grafik akurasi seperti pada gambar 8.



Gambar 8. Akurasi terbaik DNN 5 layer 450 epoch

Grafik pada gambar 8 menunjukkan grafik akurasi dari validasi dari DNN yang menggunakan 5 hidden layer selama 450 epoch. Dimana grafik yang dihasilkan menghasilkan akurasi sebesar 98.73%, presisi 99.21%, sensitivitas 99.11% dan spesifisitas 98.36%.

4. KESIMPULAN

Berdasarkan penelitian yang telah di lakukan dapat disimpulkan bahwa model terbaik dari seluruh validasi pengujian pada epoch ke 450 dengan DNN 5 hidden layer yang mampu menghasilkan akurasi sebesar 98.73%, presisi 99.21%, sensitivitas sebesar 99.11% dan spesifisitas sebesar 98.36%. Diharapkan penelitian selanjutnya mampu menguji dengan dataset jantung yang lebih banyak dan terbaru sehingga bisa menjadi acuan penelitian kedepan.

REFERENCES

- [1] A. Darmawahyuni, S. Nurmaini, and F. Firdaus, "Coronary Heart Disease Interpretation Based on Deep Neural Network," *Comput. Eng. Appl. J.*, vol. 8, no. 1, pp. 1–12, 2019.
- [2] G. S. M. Tech, "Decision Support in Heart Disease Prediction System using Naive Bayes," vol. 2, no. 2, pp. 170–176, 2011.
- [3] P. Gayathri and N. Jaisankar, "Comprehensive study of heart disease diagnosis using data mining and soft computing techniques," *Int. J. Eng. Technol.*, vol. 5, no. 3, pp. 2947–2958, 2013.



- [4] Shinde, "Heart Disease Prediction System using Multilayered Feed Forward Neural Network and Back Propagation," pp. 1–2, 2017.
- [5] D. R. Chowdhury, M. Chatterjee, and R. K. Samanta, "An Artificial Neural Network Model for Neonatal Disease Diagnosis," *Int. J. Artif. Intell. Expert Syst.*, vol. 2, no. 3, pp. 96–106, 2011.
- [6] J. Nahar, T. Imam, K. S. Tickle, and Y. P. P. Chen, "Computational intelligence for heart disease diagnosis: A medical knowledge driven approach," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 96–104, 2013.
- [7] E. O. Olaniyi, O. K. Oyedotun, A. Helwan, and K. Adnan, "Neural network diagnosis of heart disease," *2015 Int. Conf. Adv. Biomed. Eng. ICABME 2015*, pp. 21–24, 2015.
- [8] K. H. Miao and J. H. Miao, "Coronary heart disease diagnosis using deep neural networks," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 10, pp. 1–8, 2018.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2015 Inter, pp. 1026–1034, 2015.
- [10] N. A. Sundar, P. P. Latha, and M. R. Chandra, "Performance Analysis of Classification Data Mining Techniques Over Heart Disease Data Base," *Int. J. Eng. Sci. Adv. Technol.*, vol. 2, no. 3, pp. 470–478, 2012.
- [11] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, "An improved method to construct basic probability assignment based on the confusion matrix for classification problem," *Inf. Sci. (Ny)*, vol. 340–341, pp. 250–261, 2016.
- [12] G. Dahl, T. Sainath, and G. Hinton, "Improving Deep Neural Networks for LVCSR Using Rectified Linear Units and Dropout, Department of Computer Science, University of Toronto," *Acoust. Speech Signal Process. (ICASSP), 2013 IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 8609–8613, 2013.
- [13] S. K. Kumar, "On weight initialization in deep neural networks," pp. 1–9, 2017.
- [14] Y. Kwon, Y. Kwon, D. Chung, and M. Lim, "Initialization Weight Malware," no. 4, pp. 57–62, 2019.
- [15] J. jiao Xie *et al.*, "The Th17/Treg functional imbalance during atherogenesis in ApoE^{-/-} mice," *Cytokine*, vol. 49, no. 2, pp. 185–193, 2010.