






Article

Cardiovascular Disease Risk Stratification Using Hybrid Deep Learning Paradigm: First of Its Kind on Canadian Trial Data

Mrinalini Bhagawati ¹, Sudip Paul ¹, Laura Mantella ², Amer M. Johri ³, Siddharth Gupta ⁴, John R. Laird ⁵, Inder M. Singh ⁶, Narendra N. Khanna ⁷, Mustafa Al-Maini ⁸, Esmā R. Isenovic ⁹, Ekta Tiwari ¹⁰, Rajesh Singh ¹¹, Andrew Nicolaidēs ¹², Luca Saba ¹³, Vinod Anand ⁶ and Jasjit S. Suri ^{6,14,15,16,17,*}

- ¹ Department of Biomedical Engineering, North-Eastern Hill University, Shillong 793022, India; bhagawatimrinalini07@gmail.com (M.B.); sudip.paul.bhu@gmail.com (S.P.)
- ² Division of Cardiology, Department of Medicine, University of Toronto, Toronto, ON M5S 1A1, Canada; lauraevemantella@gmail.com
- ³ Division of Cardiology, Department of Medicine, Queen's University, Kingston, ON K7L 3N6, Canada; amerjohri@gmail.com
- ⁴ Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, New Delhi 110063, India; siddgupta462@gmail.com
- ⁵ Heart and Vascular Institute, Adventist Health St. Helena, St. Helena, CA 94574, USA; lairdjr@ah.org
- ⁶ Stroke Diagnostic and Monitoring Division, AtheroPoint™, Roseville, CA 95661, USA; drindersingh1@gmail.com (I.M.S.); vinodanand2010@gmail.com (V.A.)
- ⁷ Cardiology Department, Apollo Hospitals, New Delhi 110076, India; drnkhanna@gmail.com
- ⁸ Allergy, Clinical Immunology and Rheumatology Institute, Toronto, ON M5G 1N8, Canada; almaini@hotmail.com
- ⁹ Department of Radiobiology and Molecular Genetics, National Institute of The Republic of Serbia, University of Belgrade, 11001 Belgrade, Serbia; isenovic@yahoo.com
- ¹⁰ Department of Computer Science, Visvesvaraya National Institute of Technology (VNIT), Nagpur 440010, India; ekta.tiwari.ai@gmail.com
- ¹¹ Division of Research and Innovation, UTI, Uttarakhand University, Dehradun 248007, India; director.ri@uumail.in
- ¹² Vascular Screening and Diagnostic Centre, University of Nicosia, Nicosia 2417, Cyprus; anicolaides1@gmail.com
- ¹³ Department of Radiology, Azienda Ospedaliero Universitaria, 40138 Cagliari, Italy; lucasabamd@gmail.com
- ¹⁴ Department of CE, Graphic Era Deemed to be University, Dehradun 248002, India
- ¹⁵ Department of ECE, Idaho State University, Pocatello, ID 83209, USA
- ¹⁶ University Center for Research & Development, Chandigarh University, Mohali 140413, India
- ¹⁷ Symbiosis Institute of Technology, Nagpur Campus, Symbiosis International (Deemed University), Pune 412115, India
- * Correspondence: jasjit.suri@atheropoint.com; Tel.: +1-(916)-749-5628



Citation: Bhagawati, M.; Paul, S.; Mantella, L.; Johri, A.M.; Gupta, S.; Laird, J.R.; Singh, I.M.; Khanna, N.N.; Al-Maini, M.; Isenovic, E.R.; et al. Cardiovascular Disease Risk Stratification Using Hybrid Deep Learning Paradigm: First of Its Kind on Canadian Trial Data. *Diagnostics* **2024**, *14*, 1894. <https://doi.org/10.3390/diagnostics14171894>

Academic Editor: Jae-Ho Han

Received: 10 July 2024

Revised: 12 August 2024

Accepted: 26 August 2024

Published: 28 August 2024



Copyright: © 2024 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/>).

Abstract: Background: The risk of cardiovascular disease (CVD) has traditionally been predicted via the assessment of carotid plaques. In the proposed study, AtheroEdge™ 3.0_{HDL} (AtheroPoint™, Roseville, CA, USA) was designed to demonstrate how well the features obtained from carotid plaques determine the risk of CVD. We hypothesize that hybrid deep learning (HDL) will outperform unidirectional deep learning, bidirectional deep learning, and machine learning (ML) paradigms. Methodology: 500 people who had undergone targeted carotid B-mode ultrasonography and coronary angiography were included in the proposed study. ML feature selection was carried out using three different methods, namely principal component analysis (PCA) pooling, the chi-square test (CST), and the random forest regression (RFR) test. The unidirectional and bidirectional deep learning models were trained, and then six types of novel HDL-based models were designed for CVD risk stratification. The AtheroEdge™ 3.0_{HDL} was scientifically validated using *seen* and *unseen* datasets while the reliability and statistical tests were conducted using CST along with *p*-value significance. The performance of AtheroEdge™ 3.0_{HDL} was evaluated by measuring the *p*-value and area-under-the-curve for both *seen* and *unseen* data. Results: The HDL system showed an improvement of 30.20% (0.954 vs. 0.702) over the ML system using the *seen* datasets. The ML feature extraction analysis showed 70% of common features among all three methods. The generalization of AtheroEdge™ 3.0_{HDL} showed less than 1% (*p*-value < 0.001) difference between *seen* and *unseen* data, complying

with regulatory standards. Conclusions: The hypothesis for AtheroEdge™ 3.0_{HDL} was scientifically validated, and the model was tested for reliability and stability and is further adaptable clinically.

Keywords: cardiovascular disease risk; machine learning feature extraction; hybrid deep learning; scientific validation; performance evaluation; reliability; and stability

1. Introduction

Globally, the primary cause of death is cardiovascular disease (CVD) [1]. Comorbid conditions such as diabetes, hypertension, and a sedentary lifestyle contribute to the rise in CVD mortality [2,3]. Therefore, timely and non-invasive CVD risk stratification is desperately needed. A useful screening technique for the identification of cardiovascular disease and cardiovascular events (CVE) is the use of carotid ultrasonography as a surrogate biomarker [4,5]. This non-invasive diagnostic tool is highly effective in evaluating atherosclerotic plaque [6–8]. The primary indicators obtained from carotid artery ultrasound scans are carotid intima-media thickness (cIMT) [9–11], total plaque area (TPA) [11–14], and maximum plaque height (MPH) [15].

It has recently been demonstrated that a significant and trustworthy indicator for CVE and CVD is intraplaque neovascularization (IPN), an indicator of the instability in plaque and its development [16]. The CAD and carotid ultrasound-based image phenotype (CUSIP) connection were previously studied using standard regression algorithms, most of which were linear in form [17,18]. When addressing relationships of non-linear type among the risk factors and CAD ground truth, it was therefore an oversimplification. Regression-based methods also have the drawback of being unable to manage large and varied cohort sizes. These CVD risk models cannot, therefore, reliably forecast the occurrence of CVD events. Therefore, more sophisticated instruments that can precisely calibrate the CVD risk assessment are required [19,20].

Artificial intelligence (AI)-based methods have recently been demonstrated to outperform conventional statistical-based models [21]. AI has shown promise in several medical imaging applications, including radiology [21–25], dermatology [26–29], ophthalmology [30–35], cardiology [36–39], endocrinology [40–43], and, more recently, carotid ultrasonography [44–46] for the prediction of CVD risk. These AI systems can generate trained models in offline mode to understand the link between the training risk variables with the CVD outcomes [47]. To establish the events of CVD in an online form, the models that are trained are subsequently utilized to modify the test risk variables [27,48].

The field of medical imaging has recently been dominated by deep learning (DL) [49–51]. DL models are integrated with architectures having multiple layers for quick feature selection and have the added benefits of automated feature extractions [52,53]. Information from data sequences having both forward and backward temporal dependency is well captured via bidirectional models. Bidirectional systems are more robust than unidirectional ones because each element of an input sequence obtains signals from two main sources, namely the past and the present [54]. Additionally, by merging the unidirectional and bidirectional DL systems for the HDL architectural design, this study has produced sophisticated protocols.

Figure 1 displays the novel online system for the prediction of CVD risk, called AtheroEdge 3.0_{HDL}. The system displays the combination of IPN, picture phenotypes, and cardiovascular risk factors. This combination is fed into the trained system for the prediction of CVD risk, generating a risk prediction in a multi-risk granular way (shown in four colors). Three methods were used for the feature selection (not depicted in the figure): principal component analysis (PCA) pooling, the chi-square test (CST), and random forest regressor (RFR) for the machine learning system.

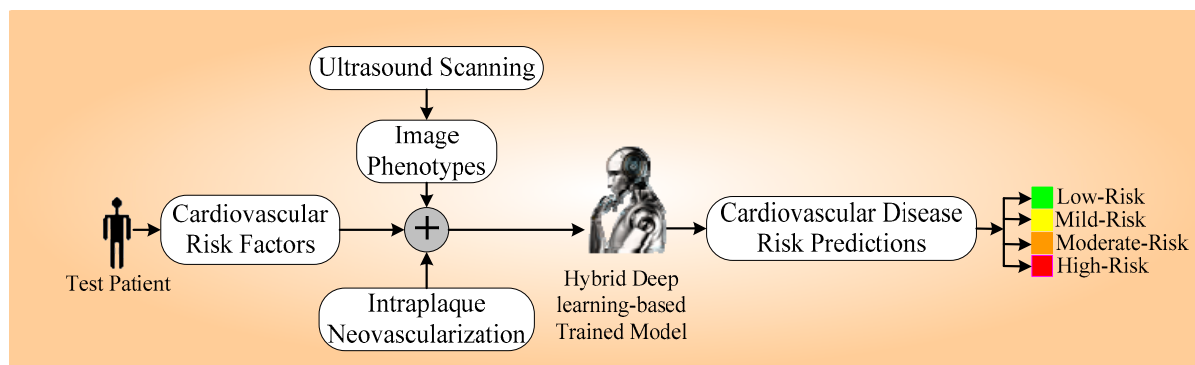


Figure 1. AtheroEdge™ 3.0_{HDL} online HDL-based system for prediction of CVD.

The AtheroEdge™ 3.0_{HDL} system undergoes performance evaluation via plotting to generate operating characteristics (ROCs) along with their p -value. The scientific validation of the AtheroEdge™ 3.0_{HDL} system was performed by using unseen data, and the reliability was checked using the statistical tests. We further benchmark the HDL-based models against unidirectional and bidirectional DL models along with the ML models (three) which are already present, namely random forest (RF) [55], gradient boost (GB) [56], and support vector machine (SVM) [56,57].

The proposed study is organized as follows: Section 2 discusses the background literature, and Section 3 contains the methodology. The results are shown in Section 4, and the performance evaluation is shown in Section 5. The statistical tests and reliability analysis are presented in Section 6. Section 7 covers the crucial discussions. Lastly, Section 8 presents conclusions.

2. The Background Literature

The evolution of CVD risk prediction systems has seen significant advancements over the years, with the integration of various methodologies. The journey started with the traditional systems between the 1980s and 1990s. Initial CVD risk prediction systems were based on traditional risk variables such as gender, age, cholesterol levels, blood pressure, and smoking status. The work by Wilson et al. [58] (1998) laid the groundwork for risk stratification in CVD and influenced subsequent studies and guidelines. The risk factor categories and the risk prediction model developed in this study have been instrumental in shaping preventive strategies and interventions for coronary heart disease (CHD).

The next phase was the inclusion of multivariate risk assessment (2000s). The focus shifted to multivariable risk assessment models that considered a broader range of factors, including diabetes, family history, and body mass index (BMI) (see Conroy et al. [59]). The systematic coronary risk evaluation (SCORE) project significantly contributed to CVD assessment in European populations. Its emphasis on region-specific models and the development of risk charts has influenced clinical practice and guidelines, providing a standardized method for evaluating CVD risk in different European contexts.

The next is the emergence of ML systems (2010), namely, RF, SVM, and neural networks, which were introduced to enhance predictive accuracy by analyzing large datasets with diverse variables. Goff Jr. et al. [60] (2014) showed that in comparison to previous risk calculators, the American College of Cardiology (ACC)/American Heart Association (AHA) calculator included a broader set of risk factors, aiming for a more comprehensive assessment of CVD. The risk calculator placed a significant emphasis on the role of statin therapy in primary prevention, recommending statin therapy for individuals with a 10-year ASCVD risk exceeding a certain threshold.

The incorporation of the image data was another advancement that included imaging phenotypes, namely cIMT (so-called, CUSIP), TPA, and coronary artery calcium scores, into risk prediction models for improved accuracy. Several scientific groups have provided examples of how to employ cIMT and TPA for CVD risk assessment [5]. AtheroEdge™

1.0 was designed which took advantage of CUSIP measurements leading to the CVD risk assessment. Subsequently, the 10-year risk prediction and CVD risk classification features of office-based biomarkers (OBBM), laboratory-based biomarkers (LBBM), and CUSIP were incorporated into the architecture of the AtheroEdge™ 2.0 system. With the AtheroEdge™ technology, scale-space image processing techniques were used for automated lumen-intima (LI) and media-adventitia (MA) border recognition for the far wall of the carotid arteries. There were several applications made available [61].

Detrano et al. [62] (2008) found a significant association between the degree to which coronary artery calcium (CAC) is present and the risk of future coronary events across all four racial or ethnic groups. The CAC scores were identified as a strong and independent predictor of coronary events, providing additional information beyond conventional risk factors. While the assessed value of CAC was consistent across the groups, some differences were observed in the CAC scores' distribution among the different racial or ethnic groups. The study highlighted the utility of CAC scoring, measured using CT scans, as a valuable tool for predicting CVD in a diverse population. This information could be valuable for risk stratification and guiding preventive interventions.

The latest and ongoing phase is DL and electronic health records (2015-present). Utilization of DL techniques includes Convolutional Neural Networks (CNNs) and recurrent neural networks (RNN), to extract patterns from electronic health records (EHRs) and improve risk prediction accuracy. Poplin et al. [63] (2018) demonstrated the feasibility of leveraging DL algorithms to extract valuable information from retinal fundus photographs for predicting cardiovascular risk factors. This innovative approach could open avenues for non-traditional methods of risk assessment in cardiovascular medicine.

3. Methodology

The role of HDL design for CVD risk stratification is the main objective. To accomplish this, we utilize the base DL models, namely unidirectional DL and bidirectional DL models. These are then combined in a tandem fashion to generate the fused HDL models. While the novel architectures are designed for CVD/stroke risk stratification, they need to be optimized via automated feature selection in the HDL system.

This section is therefore divided into the following seven subsections: The patient demographics and baseline characteristics are presented in Section 3.1. The acquisition of ultrasound images is shown in Section 3.2. The gold standard and CVD endpoint such as angiographic score is discussed in Section 3.3. The overall architecture is shown in Section 3.4, while the experimental protocol is presented in Section 3.5. The loss function adopted is shown in Section 3.6 and, lastly, Section 3.7 shows the performance metrics for AtheroEdge™ 3.0_{HDL}.

3.1. Patient Demographics and Baseline Characteristics

In total, 500 patients present in our cohort were divided into four classes based on their coronary angiography number (CAS): class 0 represented no risk or very little risk, class 1 represented mild risk, class 2 represented a moderate form of risk, and lastly, class 3 represented high risk. A total of 160 patients had acute coronary syndrome (ACS), 13 had unstable angina (UA), and 114 patients had stable angina. Moreover, 139 of the 500 patients received stents. Every subject had 39 risk variables which are further grouped into four subgroups: (i) the OBBM cluster comprising 17 covariates; (ii) the LBBM cluster with seven risk factors; (iii) the CUSIP cluster with three variables, so-called radiomics-based biomarkers; and (iv) MedUse. The OBBM cluster had the following mean scores: age (64.49 years), gender (349, 69.8%), obesity (215, 43.0%), ethnicity (486, 97.2%), BMI (31.12 kg/m²), hypertension (338, 67.6%), angina (124, 24.8%), diastolic blood pressure (76.7 mmHg), systolic blood pressure (135.35 mmHg), smoking history (330, 66%), casual-smoker (15, 3%), current-smoker (100, 20%), previous-smoker (218, 43.6%), drinks (4.94 ± 10.4 per week), family history of diabetics (195, 39.0%), premature CVD in the family (146, 29.2%), and CVD in the family (321, 64.2%). The LBBM cluster had the following

risk factors with their respective means in parentheses: creatinine ($83.99 \pm 22.6 \mu\text{mol/L}$), pre-diabetic (20, 40%), hyperlipidemia (288, 57.6%), type II diabetes (114, 22.8%), type I diabetes (5, 1.0%), estimated glomerular filtration rate ($78.96 \text{ mL/min/1.73 m}^2$), and diabetes of any type (1188, 23.6%). The CUSIP subgroup has the following means: TPA (47.68 mm^2), MPH (2.64 mm), and IPN (1.16). The MedUSE group had the following means: angiotensin-converting enzyme (ACE) inhibitors (191, 38.2%), HMG-Co reductase inhibitors (272, 54.4%), angiotensin receptor blockers (ARBs) (45, 9.0%), other antilipemic agents (9, 1.8%), alpha-blockers (30, 6.0%), calcium channel-blockers (93, 18.6%), beta-blockers (236, 47.2%), diuretics (99, 19.8%), anti-platelet medications (368, 73.6%), insulin (38, 7.6%), anti-anginals and NSAIDs (81, 16.2%), and non-insulin diabetes medications (72 (14.4%)) (see Appendix A, Table A1).

3.2. Acquisition of Ultrasound Scans and Intraplaque Neovascularization

Automated angiography screening has been popular for obtaining ultrasound scans [64,65]. It has been well proven that carotid artery disease is the surrogate biomarker for coronary artery disease [4]. Further, ultrasound imaging for the carotid artery is non-invasive and ergonomic [66,67]. All patients are made to undergo B-mode focused carotid ultrasonography imaging by using a Vivid E9 from GE Healthcare. The system has an array of 9L-D linear transducers of 2.4–10 MHz. As shown in the other studies from our group, the collection of two image phenotypes, namely TPA and MPH, was computed using longitudinal ultrasound scans using the guidelines of the American Society of Echocardiography (ASE). The MPH is the separation of the media-adventitia (MA) and the lumen-intima (LI), considering both sides of the neck. When calculating the TPA, which is the total amount of plaque present in the carotid arteries internally and externally, located 10 mm (1 cm) closest to the bifurcation of the carotid artery. According to the standard definition, when the focal structure encroaches into the lumen zone and is more than 1.5 mm in diameter or 50% of the cMT, it is classified as a carotid plaque.

The ultrasound imaging (contrast-based) of carotid artery plaque is used by the authors to quantify IPN for the detection of plaque instability and progression. IPN was graded 0, 1, 2, and 3 as per the contrasting microbubbles migration from the adventitial wall to the plaque core. These classifications are as follows: the numbers 0 and 1 indicate that there are no microbubbles visible, and 2 and 3 indicate that the plaque region is partially filled or filled with microbubbles. The total score was calculated by averaging the IPN grades on the sides of the neck. The resulting values were useful and efficient metrics for CVE and CAD.

3.3. Cardiovascular Disease Endpoint: Angiographic Score

Angiograms, graded by investigators who were blind to the clinical variables of the subjects, serve as surrogate biomarkers for CVD events. As we mentioned in our earlier work, we employ the GE Healthcare Vivid platform 2000, for the Standard Judkins ultrasound examination procedure. The group of expert cardiologists objectively marked the obtained angiograms from the coronary artery. The left anterior descending (LAD), left main, circumflex, and right coronary arteries (RCA) are the places from where the arterial stenosis grades were obtained. A disease was classified as minor if its stenosis was less than 19% in any segment, mild if it was between 20 and 49%, moderate if it was between 50 and 69%, and severe if it was greater than 70% in any segment and less than 50% in the left main coronary artery (LCA). While left ventricular echocardiography pictures can be used to automatically classify patients with CAD, coronary artery angiography is still the ground truth for CVD screening and CAD assessment. The use of a multiclass ground truth for CVD risk prediction has been demonstrated in a study employing the AtheroEdge™ 3.0_{HDL} system [54].

3.4. Overall Architecture

The architecture for AtheroEdge™ 3.0_{HDL} was designed with several DL architectures and protocols. The data are separated into subsets, namely training and testing data, where the training data are used for training the DL algorithm while testing data are used for the prediction of CVD risk. One of the key components to consider is data preparation and selecting the correct cross-validation methodology to help validate the robustness of the AI algorithm. This is accomplished by systematically partitioning the data into various non-overlapping training and testing subsets. This maximizes coverage of the learning model over all participants.

3.4.1. Data Preparation and Pre-Processing

In light of this, we divided the AtheroEdge™ 3.0_{HDL} general architecture into four main parts. The first component is for data pre-processing, which works in tandem with the second component, for data partitioning. Component three generates training models offline (see Figure 2), while component four estimates the risk of CAD or CVD using the testing datasets. Three primary procedures are involved in data preparation: (i) normalizing the data using a conventional scalar platform that places the features on a scale from 0 to 1 [68] and (ii) choosing the dominating features using three paradigms, namely PCA, CST, and RFR. The data are augmented using the synthetic minority over-sampling technique (SMOTE) [69–73] to overcome the problems of overfitting due to the small data size. Additionally, we used random sampling to further increase the data size to 5000 entries.

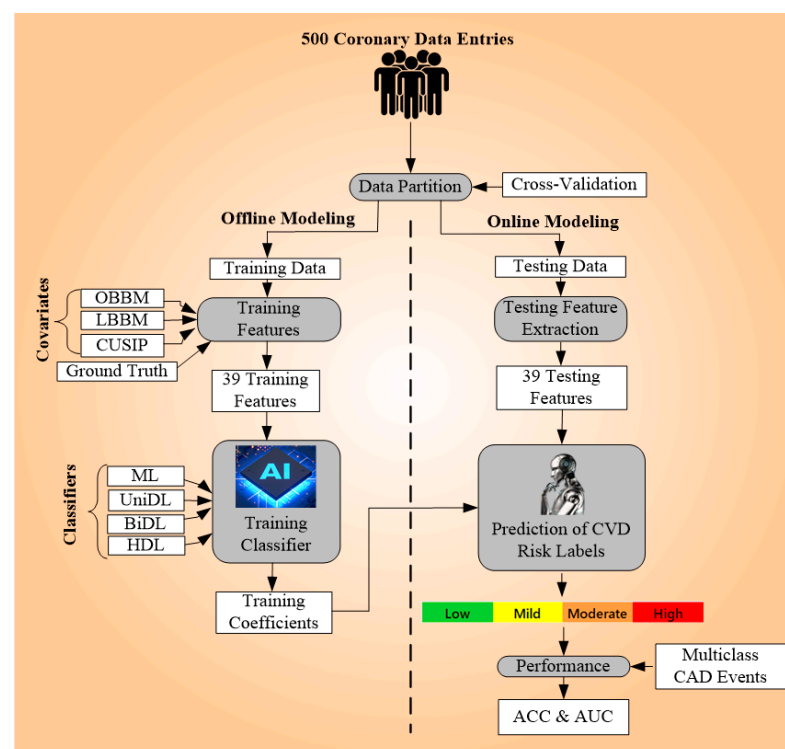


Figure 2. Overall architecture of AtheroEdge™ 3.0_{HDL}.

The data were preprocessed before being split into training and testing sets. Using the Sklearn library in Python, the quality control procedure involves applying normalization, scaling, min–max, and scaler. Each feature is normalized using the MinMaxScaler() method, which scales the data to a range. With a default of 0 and 1, the MinMaxScaler() function scales each feature separately to give the values a specified minimum and maximum value [74,75]. Unlike the self-capable DL system, the procedures for feature extraction and choosing features were solely implemented in the ML system.

3.4.2. Model Building

The third architectural component is a model-building process that uses HDL classifiers to produce the offline coefficients. Risk factors and other inputs are used to feed these classifiers. The fourth and last component is a prediction paradigm that modifies the test data to forecast the CVD risk using a trained model. The employed prototype will also predict the projected CVD risk for every ten combinations in a cyclic sequence, guaranteeing that there is no repetition of training data in the test data and that all combinations are mutually exclusive. The performance evaluation component is computed given the ground truth scores and the predicted CVD risk using the online system. In this component, we compute the AUC using ROC analysis.

3.4.3. The Rationale behind Using RNN, GRU, and LSTM

Atherosclerosis is a condition that affects people of many ethnicities and is characterized by the buildup of lipids, cholesterol, and plaque on the artery walls, which block the arteries. This study's dataset consists of a group of individuals with coronary artery disease (CAD) who are older, of the same ethnicity (usually Caucasian), and share one or more of the other three criteria. The cases under consideration share many similarities, with very slight variations. These patients' medical conditions determine which risk groups they are exposed to. We may consequently rank the patients in ascending order of coronary artery disease risk because they are all drawn from the same distribution of statistics and show up for health exams in a prearranged window of time. Mindfully, the main premise is that all subjects are relatively similar in all other aspects; it can be assumed that this new distribution represents atherosclerosis progression in the periods of discrete time for each subject depending on the risk at different phases of the disorder [54].

Sequence models, such as RNN, are among the most well-known deep learning models for their proficiency at extracting features in a temporal manner and semantics at a high-level from discrete and sequential data. This is the main rationale for classifying the risk of developing atherosclerosis using RNN and contrasting it with conventional ML-based classification systems. For such representations, LSTM, RNN, and GRU are the best options [76–79].

3.5. Experimental Protocols

To validate the hypothesis, we have created three sets of protocols. Each of the experimental protocols follows the cross-validation approach, in which the data are divided into two parts: training and testing. Four sets of cross-validation paradigms are implemented, namely two-fold (K2), four-fold (K4), five-fold (K5), and ten-fold (K10), where K2, K4, K5, and K10 constitute the following partition ratios: 50:50, 75:25, 80:20, and 90:10. Each of the cross-validation protocols is implemented in three kinds of experimental protocols, namely experimental protocol 1 (EP1), experimental protocol 2 (EP2), and experimental protocol 3 (EP3).

3.5.1. Experimental Protocol 1: Three Unidirectional Models

The unidirectional DL models RNN [80], gated-recurrent units (GRU) [81,82], and long short-term memory (LSTM) [83] have all been taken into consideration by the authors in EP1. RNNs work particularly well in sequence-based applications such as series-of-time prediction, audio recognition, and processing of natural language. However, ordinary RNNs struggle to capture long-term correlations due to issues like bursting or fading gradients (see Appendix B, Figure A1). To address the vanishing gradient problem, one type of RNN architecture known as LSTM (see Appendix B, Figure A2) was developed. Its memory cells and various information-flow regulating gates make it more efficient at determining distant links within sequences. GRU is another RNN variant that is similar to LSTM but has a marginally different topology. It has been proven to perform well in a variety of sequence-based tasks and is outfitted with gating devices to control the flow of

information (see Appendix B, Figure A3). These models are utilized for K2, K4, K5, and K10, which are the various CV regimens.

3.5.2. Experimental Protocol 2: Three Bidirectional Models

The EP2 had the bidirectional DL systems, namely BiRNN [84] (See Appendix C, Figure A4), BiLSTM [85] (See Appendix C, Figure A5), and BiGRU [77,86] (See Appendix C, Figure A6).

BiRNN

In order to analyze sequential data, the input sequence traverses in two directions, namely, forward and backward, at same the time in BiRNN. The network performs a forward pass to gather data from the past and a backward pass for gathering data from the future for each time step. Therefore, the representation of the context present in each time step is holistically accomplished via the combination of the hidden states from both sides. Due to the bidirectional manner of operation, BiRNNs are mainly beneficial for applications where deep context comprehension is required, such as time-series analysis, recognition of voices, processing natural languages, CVD risk prediction, and depressive illness prediction. The network can more efficiently recognize dependencies and connections that are present in the sequential data. By using backpropagation to adjust the parameters of the network during training, the accuracy of the predictions or classifications the network may make based on the bidirectional contextual input is maximized.

BiLSTM

The BiLSTM model is based on the bidirectional flow for understanding the sequential input. It records the information in a contextual form from both the past and future by comparing the forward and backward layers of the LSTM model. By managing input, forget, and output gates, the LSTM units govern the flow of information and preserve cells of memory for storing relevant data at each time step. The neural network may comprehend long-term relationships and linkages along the whole sequence as they are of bidirectional nature. Concatenating the hidden states that are present on both sides generally results in the representation of the context more completely. This architecture is particularly useful for issues like time series, natural language processing, and illness prediction where accurate prediction necessitates knowledge of both preceding and subsequent components. Back-propagation changes the parameters of the network during training over time, which improves the capacity for representing the sequential relationship.

BiGRU

BiGRU is based on two parallel layers, namely forward and backward layers, for assessing the sequential input. Every time step, data are continuously collected by the forward GRU from earlier items and by the reverse GRU from later elements. Then, to create a complete picture of the context, the concealed states on both ends are frequently combined. BiGRU networks are useful for characterizing linkages and interactions within sequential data since they are bidirectional. Throughout the training, backpropagation processes gradually update the network's properties, optimizing the capacity of the network to generate accurate classifications based on reciprocal contextual input.

3.5.3. Experimental Protocol 3: Six Hybrid Models

EP3 features a hybrid architecture which has the combination of the unidirectional DL and bidirectional DL systems. We design a combination of unidirectional DL or bidirectional DL models. In combination A, we combine both the models to create a unidirectional DL model. In combination B, we have combined both models to be bidirectional DL models. In combination A, we have three combinations, namely RNN + GRU, RNN + LSTM, and LSTM + GRU. In combination B, there are three combinations, namely: BiRNN + BiGRU, BiRNN + BiLSTM, and BiLSTM + BiGRU (See Appendix D, Figure A7). The HDL models

were created by connecting the unidirectional and bidirectional models in tandem to obtain a better performance when compared to other models.

3.6. Loss Function and Training Parameters

The categorical cross entropy (CCE) loss function [87–89] is used for the calculation of the loss during the training of the system. Further, the losses obtained during the training and validation of the system were plotted against the number of epochs (100) present in the system. This type of loss function is used for the balanced distribution of data. The CCE loss function was selected over the other types as it is best suited for the binary variables. All other loss function types can be further added as future work. The binary cross entropy loss has been described in the equation below:

$$\mathcal{L}_{\text{CCE}} = -(g \log(p) + (1 - g) \log(1 - p)) \quad (1)$$

where \mathcal{L}_{CCE} is the binary cross-entropy loss, g represents the true label of the i th instance (either 0 or 1 for binary classification), and p represents the predicted probability that the i th instance belongs to class 1. Additionally, the training parameter used is as follows: the optimizer used is Adam [90], the number of epochs is 100, the batch size is 32, and the learning rate is 0.001. We utilized the early stopping method to reduce overfitting. Using the specified parameters, we trained all 12 models and their combinations with the 4 different cross-validation protocols. Afterward, we determined the most optimal model by comparing their performance metrics and selecting the best model based on this comparison.

3.7. Performance Metrics

The following phrases are utilized to derive the performance metrics: The number of cases that the model correctly classifies as positive is known as the True Positive (TP) value. The number of cases that the model correctly classifies as negative is known as the True Negative (TN) value. The number of times a negative instance is mistakenly classified as positive via the model is known as the False Positive (FP) value. The number of occurrences that the model mistakenly classifies as negative when they are positive is known as the False Negative (FN) value.

The percentage of accurate forecasts among all the predictions the model makes is known as accuracy. The percentage of accurate positive forecasts among all real positive occurrences is known as sensitivity. The performance indicator known as specificity counts the percentage of false positives that the model misidentified out of all the real negative cases. For varying classification thresholds, the receiver operating characteristic (ROC) curve represents the trade-offs between the TPR and FPR. By computing the area under the ROC curve, the area under the curve (AUC) measures the overall quality of the model.

$$\text{Accuracy(ACC)} = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (2)$$

$$\text{Sensitivity(SEN)} = \frac{TP}{(TP + FN)} \quad (3)$$

$$\text{Specificity(SPEC)} = \frac{FP}{(FP + TN)} \quad (4)$$

4. Results

The obtained results for accuracy and loss, corresponding to each experimental protocol, along with the comparison of three feature selections are discussed here in five subsections, namely Sections 4.1–4.5, for accuracy vs. loss curves, results of experimental protocol 1, results of experimental protocol 2, results of experimental protocol 3, and comparison of three feature selections, respectively. The obtained results concur with our hypothesis, hence proving our hypothesis to be correct.

4.1. Accuracy and Loss Curves

The model accuracy and the loss were plotted against the epochs used in the system. The accuracy of the model increased with the increasing number of epochs, whereas the loss decreased with the increasing epochs for both training and validation. The curve showed the exponential rise for the accuracy plot and the exponential decay for the loss functions. The loss function considered was categorical cross entropy. The accuracy and loss curve for the BiLSTM + BiGRU models are displayed in Figure 3 as an example of where the Adam optimizer is applied.

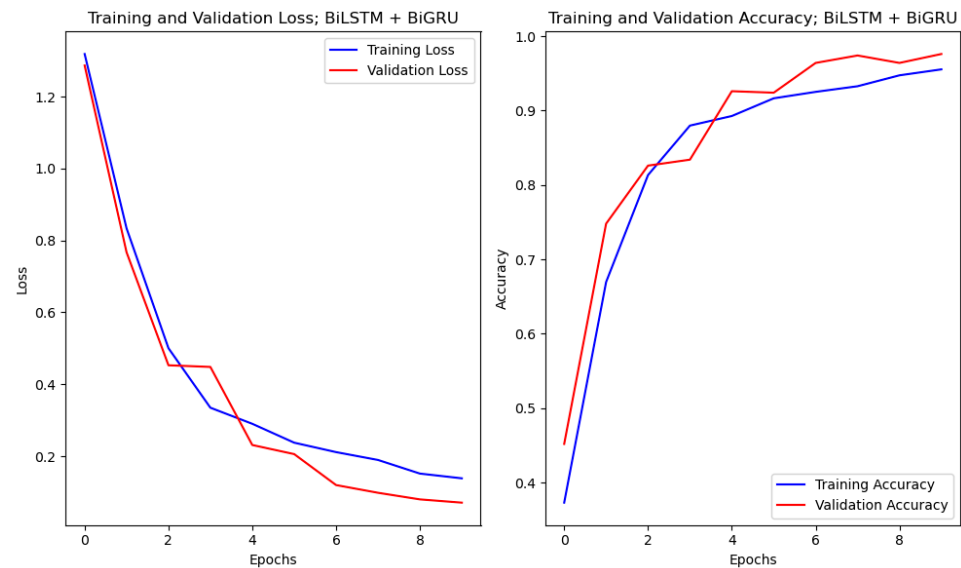


Figure 3. (Left) Loss vs. epochs plot for the BiLSTM + BiGRU model; (Right) accuracy vs. epochs plot for the BiLSTM + BiGRU model.

4.2. Results of Experimental Protocol 1

Table 1 details the results parameter for the unidirectional DL systems, namely RNN, GRU, and LSTM, for the different protocols (K2, K4, K5, and K10). The LSTM architectures have the highest accuracy and AUC in all the protocols. The K10 protocol provides the highest accuracy for all the unidirectional DL systems. The AUC ranges from 0.883 to 0.907 for all three models in the K2 protocol, 0.850 to 0.908 in the K4 protocol, 0.895 to 0.916 (RNN, GRU, and LSTM) in the K5 protocol, and 0.896 to 0.918 for all three models (RNN, GRU, and LSTM) in the K10 protocols.

Table 1. Results for all unidirectional DL models for all cross-validation protocols.

CVP	Models	ACC (%)	SPEC (%)	SEN (%)	<i>p</i> -Value	AUC (0–1)
K2	RNN	80.11	82.96	81.87	<0.001	0.883
	GRU	81.12	83.32	82.16	<0.001	0.892
	LSTM	81.81	85.67	83.36	<0.001	0.907
K4	RNN	82.52	81.97	83.88	<0.001	0.850
	GRU	83.13	82.33	84.17	<0.001	0.897
	LSTM	84.82	84.68	85.37	<0.001	0.908
K5	RNN	83.50	82.95	84.86	<0.001	0.895
	GRU	84.11	83.31	85.15	<0.001	0.900
	LSTM	85.80	85.66	86.35	<0.001	0.916
K10	RNN	84.20	83.15	83.86	<0.001	0.896
	GRU	85.12	84.11	86.15	<0.001	0.902
	LSTM	86.80	86.67	87.35	<0.001	0.918

CVP: Cross-Validation Protocol; ACC: Accuracy; SPEC: Specificity; SEN: Sensitivity; AUC: Area-Under-the-Curve.

4.3. Results of Experimental Protocol 2

The results obtained for the different bidirectional DL systems are displayed in Table 2. The different protocols used were K2, K4, K5, and K10. It can be seen that the BiLSTM architecture has the highest accuracy and AUC in all the protocols as well as, the best results in obtained in the K10 protocol for all the bidirectional DL models. The BiLSTM architecture has the ability for bidirectional mapping and enhanced contextual understanding that results in higher performance. The BiLSTM models provide flexible manipulation or exploration of the latent space.

Table 2. Results for all bidirectional DL models for all K protocols.

CVP.	Models	ACC (%)	SPEC (%)	SEN (%)	p-Value	AUC (0–1)
K2	BiRNN	82.22	80.16	82.34	<0.001	0.891
	BiGRU	83.42	81.17	83.18	<0.001	0.905
	BiLSTM	84.32	83.29	84.12	<0.001	0.913
K4	BiRNN	83.28	82.66	83.94	<0.001	0.910
	BiGRU	84.24	83.12	84.28	<0.001	0.915
	BiLSTM	85.13	85.68	86.52	<0.001	0.923
K5	BiRNN	84.28	83.66	84.94	<0.001	0.920
	BiGRU	85.24	84.12	85.28	<0.001	0.925
	BiLSTM	86.13	86.68	87.52	<0.001	0.933
K10	BiRNN	85.27	84.65	85.93	<0.001	0.920
	BiGRU	86.23	85.11	86.27	<0.001	0.925
	BiLSTM	87.00	86.67	88.51	<0.001	0.935

CVP: Cross-Validation Protocol; ACC: Accuracy; SPEC: Specificity; SEN: Sensitivity; AUC: Area-Under-the-Curve.

4.4. Results of Experimental Protocol 3

EP3 featured the HDL systems. Table 3 details the different parameters that are obtained for the 12 combined HDL systems. The result showed that the AUC ranged from 0.930 to 0.974 in the K2 protocol, 0.931 to 0.975 in the K4 protocol, 0.947 to 0.985 in the K5 protocol, and 0.948 to 0.985 in the K10 protocol, showing the best performance in the K10 protocol. Among all the HDL systems, the combination of bidirectional DL systems (BiRNN + BiGRU, BiRNN + BiLSTM, and BiLSTM + BiGRU) results in the highest AUC.

Table 3. Results for all HDL models for all K protocols.

CVP	Models	ACC (%)	SPEC (%)	SEN (%)	p-Value	AUC (0–1)
K2	RNN + GRU	85.00	82.76	82.62	<0.001	0.930
	RNN + LSTM	85.91	85.37	83.56	<0.001	0.939
	LSTM + GRU	88.02	88.16	84.62	<0.001	0.943
	BiRNN + BiGRU	90.53	89.75	88.32	<0.001	0.954
	BiRNN + BiLSTM	91.84	90.18	88.32	<0.001	0.968
	BiLSTM + BiGRU	94.14	92.22	89.12	<0.001	0.974
K4	RNN + GRU	86.66	83.76	83.62	<0.001	0.931
	RNN + LSTM	86.02	86.37	84.56	<0.001	0.940
	LSTM + GRU	89.02	89.16	85.62	<0.001	0.944

Table 3. Cont.

CVP	Models	ACC (%)	SPEC (%)	SEN (%)	p-Value	AUC (0–1)
	BiRNN + BiGRU	91.53	90.75	87.32	<0.001	0.955
	BiRNN + BiLSTM	92.84	91.18	89.32	<0.001	0.969
	BiLSTM + BiGRU	95.14	93.22	90.12	<0.001	0.975
K5	RNN + GRU	86.45	82.76	84.62	<0.001	0.947
	RNN + LSTM	86.50	85.37	86.56	<0.001	0.957
	LSTM + GRU	91.02	88.16	87.62	<0.001	0.959
	BiRNN + BiGRU	93.53	84.75	87.32	<0.001	0.972
	BiRNN + BiLSTM	94.84	86.18	87.32	<0.001	0.976
	BiLSTM + BiGRU	96.14	92.22	93.12	<0.001	0.982
K10	RNN + GRU	88.14	82.75	84.61	<0.001	0.948
	RNN + LSTM	88.24	84.36	85.55	<0.001	0.958
	LSTM + GRU	92.01	87.15	86.61	<0.001	0.960
	BiRNN + BiGRU	94.52	88.74	87.31	<0.001	0.975
	BiRNN + BiLSTM	95.83	89.17	87.31	<0.001	0.979
	BiLSTM + BiGRU	97.25	92.21	93.11	<0.001	0.985

CVP: Cross-Validation Protocol; ACC: Accuracy; SPEC: Specificity; SEN: Sensitivity; AUC: Area-Under-the-Curve.

4.5. Comparison of Three Feature Selection Methods for Three Machine Learning Models

The top important features were identified by using the three methods, namely PCA, CST, and RFR. The features selected are displayed in Table 4. The top ten commonly selected features are age, diabetes T1D, avg system before angio, IPN, creatinine, hyperlipidemia, and alpha-blockers. The mean accuracy increased with the involvement of these features. The mean accuracy of the HDL models is 0.956 with the selection of the features.

Table 4. The top ten selected features for all three methods.

SN	PCA	CST	RFR
1	Age	Age	Age
2	Diabetes T1D	Diabetes T1D	Diabetes T1 D
3	Avg Sys before angio	Avg Sys before angio	Avg Sys before angio
4	IPN	IPN	IPN
5	Creatinine	Creatinine	Creatinine
6	Hyperlipidemia	Hyperlipidemia	Hyperlipidemia
7	Alpha-Blockers	Alpha-Blockers	Alpha-Blockers

Table 4. Cont.

SN	PCA	CST	RFR
8	Insulin	Family Hx of CVD	Current Smoker
9	Angina	Current Smoker	BMI
10	Anti-Platelet/Anti-Coagulants	Anti-Platelet/Anti-Coagulants	TPA

SN: serial number; PCA: principal component analysis; CST: chi-square test; RFR: random forest regressor; IPN: intraplaque neovascularization; BMI: body mass index; TPA: total plaque area.

5. Performance Evaluation

The evaluation of any AI system must include the performance of the AI system. This includes the prediction and its comparison with gold standard labels, which can be accomplished by analyzing the receiver operating characteristics (ROC). Further, we are interested in understanding the effect of the training data size on the performance of the AI system. Lastly, we must quantify how HDL models meet the hypothesis that they are superior to unidirectional DL, bidirectional DL, and ML systems. In this context, the ROC is presented in Section 5.1, the effect of sample size is shown in Section 5.2, and the superiority of HDL models over other models is shown in Section 5.3.

5.1. Receiver Operating Characteristics of All Four Kinds of AI Models

Figure 4 shows the ROC curves obtained for all four clusters of the system, namely ML, unidirectional DL, bidirectional DL, and HDL, in increasing order of AUC values with corresponding *p*-values. The HDL cluster has the highest mean AUC value of 0.964 represented in the black color curve, whereas it can be observed that the ML system has the lowest mean AUC of 0.702 represented in the red color curve. In between these two curves lies the curve for the unidirectional DL in green color with a mean AUC of 0.910 and bidirectional DL in violet color with a mean AUC of 0.931. The fusion of unidirectional DL and bidirectional DL models demonstrates the rationale for the highest AUC validating our hypothesis. The corresponding *p*-values ensure the significance of the AUC values.

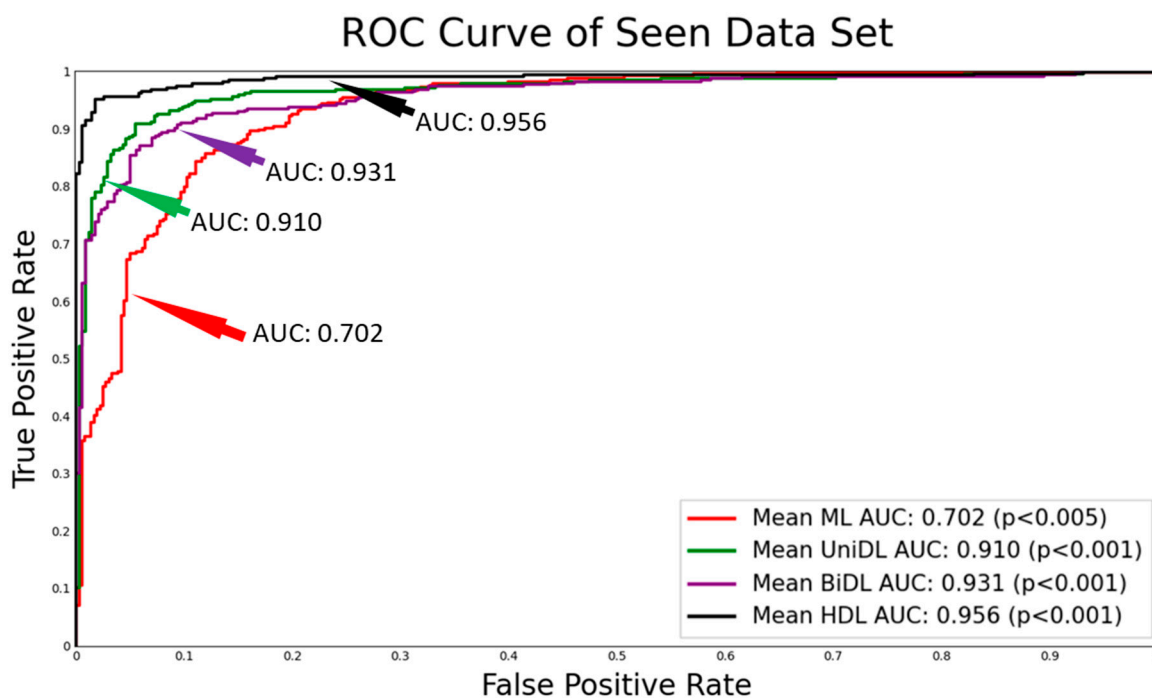


Figure 4. ROC showing the mean AUC along with their *p*-values; AUC: area-under-the-curve; ML: machine learning; DL: deep learning; UniDL: unidirectional DL; BiDL: bidirectional DL; HDL: hybrid DL.

5.2. Effect of Sample Sizes on the Training System

The idea behind this experiment is to understand the effect of training data size on the performance of the AI models. This will assure us that the training models can better generalize, unlike memorization. To accomplish this experiment, we have taken four sets of paradigms, namely K2, K4, K5, and K10, which are computed for all the AI models. To appreciate the effect of data size on the training of the models, we compute the “stacked accuracy for (SA)” by stacking the accuracies for each of the cross-validation protocols.

This stacking accuracy is computed for all four AI models. This allows for a powerful form of visual representation. Figure 5 shows the effect of training data size for all four models using the stacked accuracy concept. As seen in the figure, we can appreciate the color bands of each block gradually increasing compared to the previous color block. The best increase in color bands was seen for the HDL model, unlike the unidirectional, bidirectional, and machine-learning models. This validates our assumption that the performance improves with an increase in the training data size. This also validates that the increase in accuracy is very gradual and symptomatically settles down to nearly constant values even if the training set is 90% of the data size.

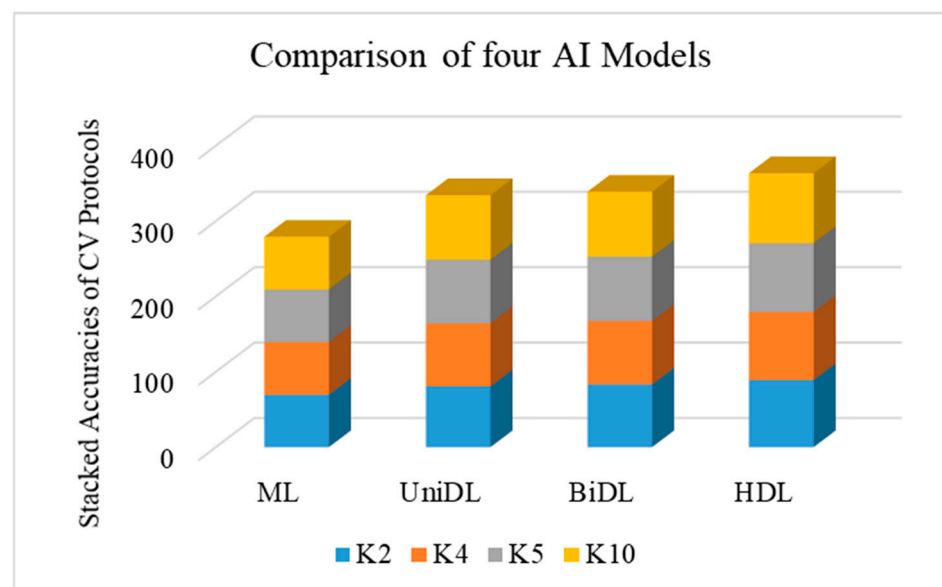


Figure 5. Plots for the effect of data size in the four model types; ML: machine learning; DL: deep learning; UniDL: unidirectional DL; BiDL: bidirectional DL; HDL: hybrid DL.

5.3. Hybrid Deep Learning Performance against Other AI Models

While HDL has shown superior performance compared to ML, unidirectional, and bidirectional models, it is important to quantify the absolute improvement in the difference between HDL and other AI models. The HDL system showed an improvement of 30.20% over the ML system using the seen datasets (since the mean accuracy of HDL was 90.88%, and the mean accuracy of ML was 69.80%; thus, the improvement was computed as the absolute difference of the mean accuracy of HDL minus the mean accuracy of ML divided by the mean accuracy of ML times 100, i.e., $|90.88 - 69.80| / 69.80 \times 100 \sim 30.2\%$). Table 5 displays the absolute percentage increase improvement in HDL models when compared to ML, unidirectional DL, and bidirectional DL models. Our observations show that the HDL model is superior to ML, unidirectional DL, and bidirectional DL models by 30.20%, 8.72%, and 7.26%, respectively.

Table 5. Comparison of ML vs. unidirectional DL vs. bidirectional DL vs. HDL.

SN	M1 (a)	M2 (b)	% Increase $(a - b/a) \times 100$
1	ML	HDL	30.20%
2	UniDL	HDL	8.72%
3	BiDL	HDL	7.26%

ML: machine learning; DL: deep learning; UniDL: unidirectional DL; BiDL: bidirectional DL; HDL: hybrid DL.

6. Scientific Validation

The best way to validate the AI system is to evaluate the AI on a dataset that was never part of the original datasets [91,92]. This means the training was conducted on dataset A and testing on dataset B, where dataset B was never part of data A (See Appendix E, Table A2) [93]. To conduct the scientific validation, we took dataset A as the experimental dataset and dataset B as the validation dataset from another source having 20 risk predictors. Thus, our validation dataset consisted of 459 subjects with 20 risk features. Table 6 shows side by side the mean AUC values (along with their *p*-values) for experimental data (seen dataset) and validation data (unseen data). If one can show that the difference between these results has a 5% range, the AI model can be characterized for regulatory compliance. The percentages difference among the *seen* and *unseen* data were 2.78%, 2.94%, 2.87%, and 1.79% for ML, unidirectional DL, bidirectional DL, and HDL systems, respectively. We did observe that HDL models showed the lowest absolute difference between the experimental data (see analysis) and validation data (unseen data analysis) (see Table 6, Figure 6).

Table 6. Mean AUC values for seen vs. unseen data.

SN	Models	Seen Data		Unseen Data		
		Mean AUC (0–1)	<i>p</i> -Value	Mean AUC (0–1)	<i>p</i> -Value	Difference (%)
1	ML	0.702	<0.005	0.683	<0.005	2.78%
2	UniDL	0.910	<0.001	0.884	<0.001	2.94%
3	BiDL	0.931	<0.001	0.905	<0.001	2.87%
4	HDL	0.956	<0.001	0.939	<0.001	1.79%

ML: machine learning; DL: deep learning; UniDL: unidirectional DL; BiDL: bidirectional DL; HDL: hybrid DL.

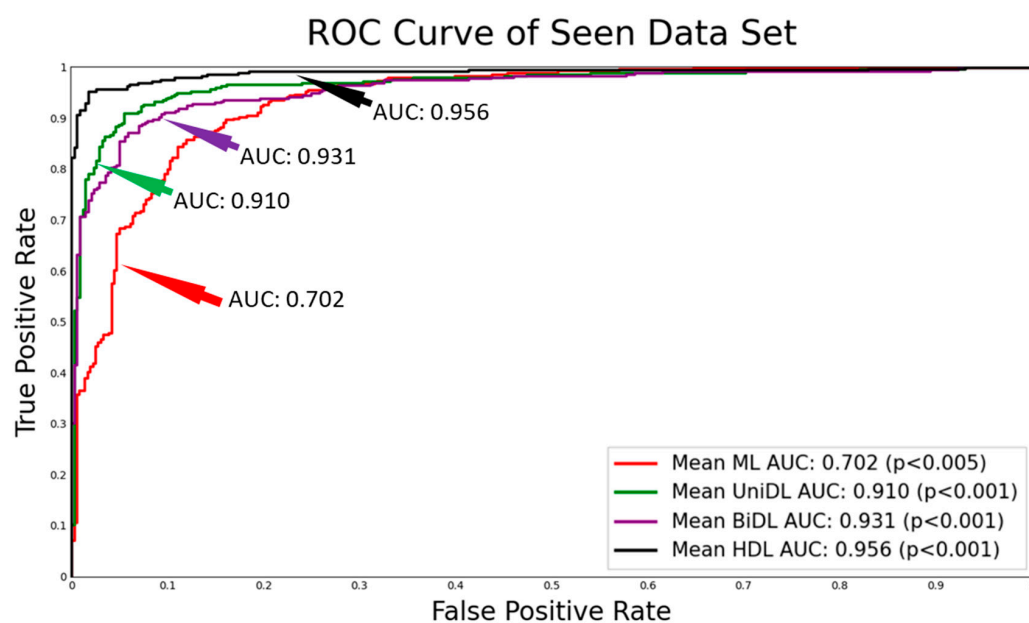


Figure 6. Cont.

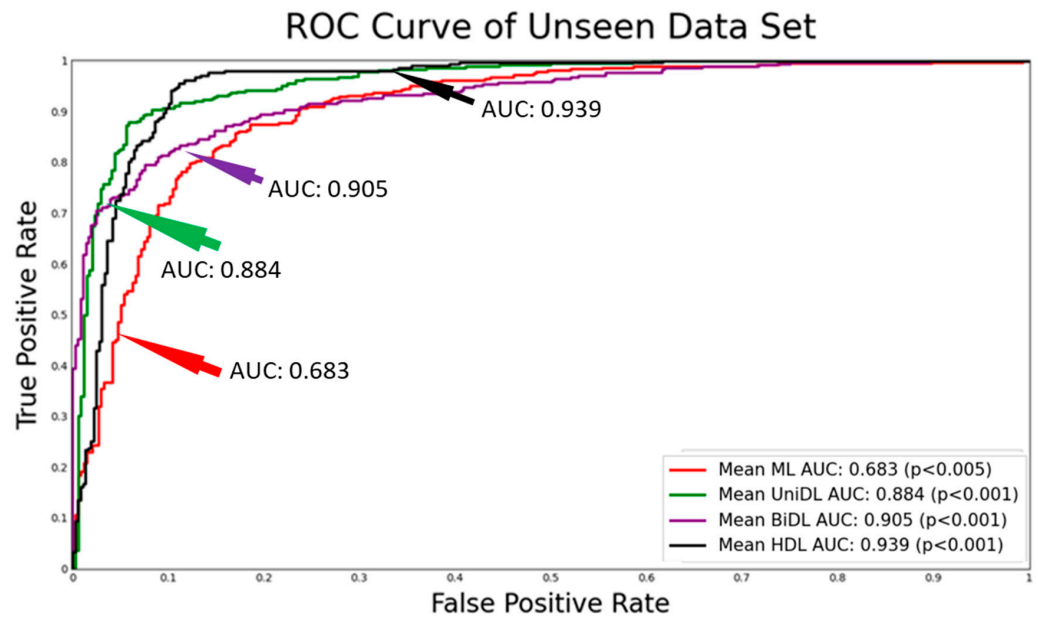


Figure 6. Receiver operating characteristic curve for mean AUC; (**Top**): seen dataset; (**Bottom**): unseen dataset.

7. Reliability and Stability of Hybrid Deep Learning System

The reliability analysis was performed for the unidirectional (RNN) vs. bidirectional (BiRNN), HDL model (BiGRU + BiRNN) vs. unidirectional model (RNN), and HDL model (BiGRU + BiRNN) vs. bidirectional model (BiRNN), and more, with the K10 protocol by using statistical tests, namely the Mann–Whitney test. A non-parametric statistical test called the Mann–Whitney U test is used to see if there is a difference between two independent groups. It can be used for ranking, summing ranks, test statistics, and comparing critical values. It is specifically designed for paired data or repeated measures, where each subject or entity is measured twice under different conditions. The *p*-values were calculated and are displayed in Table 7. The *p*-values obtained are very similar, being <0.05 in all three compared models, hence validating that the systems are reliable and stable. A thorough selection of models utilizing empirical data is necessary, as indicated by the consistently low *p*-values, which show significant and persistent changes in prediction accuracy.

Table 7. Reliability and stability results for best unidirectional vs. bidirectional hybrid DL models.

SN	Model1	Model2	Mann-Whitney
1	BiRNN	RNN	<i>p</i> < 0.05
2	BiGRU + BiRNN	RNN	<i>p</i> < 0.05
3	BiRNN + BiLSTM	BiRNN	<i>p</i> < 0.05
4	BiRNN + BiLSTM	RNN	<i>p</i> < 0.05
5	BiRNN + BiGRU	BiRNN	<i>p</i> < 0.05
6	BiRNN + BiGRU	RNN	<i>p</i> < 0.05
7	BiLSTM + BiGRU	RNN	<i>p</i> < 0.05
8	BiLSTM + BiGRU	BiRNN	<i>p</i> < 0.05

8. Discussion

The proposed novel study is unique in its application of HDL mechanisms for CVD risk stratification. Here, we have combined the angiographic scores (gold standard) with the risk variables such as CUSIP, IPN, and the conventional clinical risk factors. Leveraging the AtheroEdge™ 3.0_{HDL} architecture, this is the first study where the combination of traditional risk factors, laboratory-based factors, image phenotypes (radiomic features),

and MedUSE risk factors was implemented in an HDL framework and benchmarked against machine learning, unidirectional DL, and bidirectional DL frameworks. The study demonstrated six HDL models. In combination A, we have three combinations, namely RNN + GRU, RNN + LSTM, and LSTM + GRU. In combination B, there are three combinations, namely BiRNN + BiGRU, BiRNN + BiLSTM, and BiLSTM + BiGRU. For input into the AtheroEdge™ 3.0_{HDL} architecture, the top 10 important features were identified via three methods, namely PCA, chi-square, and RFR. All three methods showed that 70% of features were shared amongst these three methods. The HDL system, AtheroEdge™ 3.0_{HDL}, was more accurate for risk stratification of CVD than bidirectional DL, unidirectional DL, and ML models. Further, our study showed a ~30.20% improvement in the AUC for predicting the CVD in the HDL framework compared to the ML strategy. A statistical test, namely the Mann–Whitney test, was carried out to test the statistical reliability of the HDL AI system. For generalization of the AtheroEdge™ 3.0_{HDL}, the training was performed using experimental datasets, and testing was performed on the validation datasets. Such a system showed less than 1% difference between the seen data analysis and unseen data analysis, which also satisfies regulatory compliance, where the requirement is less than 5% [94,95].

8.1. Benchmarking Table

The benchmarking table for CVD risk stratification, which includes eight DL-based studies, is shown in Table 8. The characteristics include the author, the total features used (NOF), the strategies used, the total number of subjects/images, the cross-validation (CV) methodology, the survival analysis (SA), and the outcomes.

Table 8. Benchmarking.

C0	C1	C2	C3	C4	C5	C6	C7
SN	Authors	NOF	ML/DL Models Used	#Patients/#Ima	CV	SV	Results Obtained
R1	Unnikrishnan et al. [96]	09	SVM	2.4 K	K5	×	AUC for ML = 0.71; for CCVRC = 0.57
R2	Jamthikar et al. [56]	39	RF, SVM, XGBoost	500	K10	✓	AUC for ML = 0.95; for CCVRC = 0.50
R3	Alaa et al. [19]	473	SVM, GBM, RF, AdaBoost	423.6 K	K10	×	AUC for ML = 0.724; for CCVRC = 0.774
R4	Zhou et al. [97]	×	UNet++	144/510 497/638	K5	×	TPA error = 5.55 ± 4.34 mm ² ; DSC = 83.3–85.7%
R5	Jain et al. [98]	×	UNet+, UNet, SegNetUNet, SegNet, Unet + SegNet	97/970	K5	×	AUC for UNet = 0.91, for UNet + = 0.91, for SegNet-UNet = 0.908, for SegNet = 0.905, and for SegNetUNet+ = 0.898 (using CE-loss models) and 0.883, 0.889, 0.905, 0.889, and 0.907 (using DSC-loss models); PA error = 3.49 mm ² for SDL and 4.21 mm ² for HDL
R6	Jain et al. [99]	24	UNet	379, 300	K10	×	Unseen FoM: 70.96 and 91.14 Seen FoM: 97.57, 88.89, and 99.14

Table 8. Cont.

C0	C1	C2	C3	C4	C5	C6	C7
SN	Authors	NOF	ML/DL Models Used	#Patients/#Ima	CV	SV	Results Obtained
R7	Jain et al. [100]	24	× AtheroEdge 2.0, UNet, UNetSegNet	379	K10	×	AUC for UNet = 0.93, for SegNet-UNet = 0.94; for AtheroEdge™ 2.0 = 0.95, respectively; SDL PA error = 9.9 mm ² ; HDL PA error = 8 mm ² ; AtheroEdge™ = 2.0 mm ² ; PA error = 9.6 mm ²
R8	Johri et al. [54]	39	RF, SVM, RNN, LSTM	500	K10	✓	AUC for DL AUC = 0.99, for ML = 0.89, for CCVRC = 0.50
R9	Akari et al. [48]	56	SVM	4004	K10	✓	ACC = 98.43; reliability index = 97.32%
R10	Proposed method	39	SVM, RF, XGBoost, UniDL, BiDL, HDL	500	K2, K4, K5, K10	✓	AUC for ML = 0.702, for UniDL = 0.910, for BiDL = 0.931, and for HDL = 0.956

SN: Serial number; NoF: #features; DL: Deep learning; ML: Machine learning; RF: Random forest model; SVM: Support vector machine model; GBM: Gradient boost machine model; CV: Cross-validation; SV: Scientific validation; RNN: Recurrent neural network; CCVRC: Conventional cardiovascular risk calculator; LSTM: Long short-term memory; AUC: Area-under-the-curve; SDL: Solo DL; TPA: Total plaque area; PA: Plaque area; HDL: Hybrid DL; ACC: Accuracy; uniDL: Unidirectional DL; BiDL: Bidirectional DL; HDL: Hybrid DL.

Nine risk variables were present in each of the 2406 patients in the Australian datasets utilized in the first analysis by Unnikrishnan et al. [96] (R1). They have chosen to use the conventional cardiovascular risk calculator (CCVRC) and linear regression classifiers in the K5 protocol for the Framingham Risk Score (FRS). The authors demonstrated that, in contrast to the ML-based calculators' 0.71 AUC, the 10-year risk assessment's AUC for CCVRC was 0.57. Next, an investigation by Jamthikar et al. [56] (R2) compared ML calculators, such as RF, GB, and SVM, with conventional calculators, such as the systematic coronary risk evaluation (SCORE), FRS, and extreme atherosclerotic cardiovascular disease (ASCVD). A total of 500 participants with 39 risk variables were used in this investigation. The authors demonstrated a significant difference in AUC between the CCVRC and ML models, with *p*-values less than 0.0001 and AUC values of 0.50 and 0.95, respectively. In the third study, which was conducted by Alaa et al. [19] (R3), the authors used a five-year follow-up technique to show how well ML classifiers performed versus the FRS algorithm for a UK dataset containing 423,604 subjects with 473 variables. As compared to CCVRC, the outcomes for ML classifiers were more accurate. The ML classifiers were AdaBoost, RF, SVM, and gradient-boosted machine (GBM) models with an AUC equal to 0.774 for the ML and 0.724 for the CCVRC.

The use of carotid imaging for CVD risk stratification has been increasing. Zhou et al., the fourth study (R4) [97], demonstrated the use of the UNet++ model for plaque segmentation in a multi-ethnic database. Depending on the CV methodology that was employed, the training and testing datasets included 100 and 44 images, respectively. Due to the small amount of data, the system was unable to be adequately justified for the hospital settings. They did not even benchmark the system against other cutting-edge systems. The HDL models (SegNetUNet and SegNetUNet+) were presented for segmentation of the plaque in a different work by Jain et al. [98] (R5). The internal carotid artery (ICA) was the section of the artery under consideration. To improve the datasets, the rotation transformation augmentation technique was used. The method's shortcomings included racial bias in data selection and source identification. Two sets of multi-ethnic CCA datasets from Japan and Hong Kong were employed by Jain et al. [99] (R6) in their

tests. To prevent various biases, they suggested using the HDL architecture for the unseen dataset. Nevertheless, the authors neglected to validate the suggested method against any existing CVD risk prediction systems on the market, which resulted in bias in the system's validation. The seventh study taken into consideration was authored by Jain et al. [100] (R7), in which the solo DL (SDL) models and HDL models are contrasted with the different available systems in the market (AtheroEdge™ 2.0) suggested by AtheroPoint™ LLC, CA, USA. TPA inaccuracies for HDL (8 mm²), SDL (9.9 mm²), and traditional models (9.6 mm²) were displayed in the results for the image datasets. Johri et al. [54] (R8) show the use of 39 features with the ML models, namely RF and SVM. The used DL models were RNN and LSTM, applied to data from a total of 500 patients by following the K10 protocol. The results obtained were an AUC for the DL model of 0.99, for ML models 0.89, and for the CCVRC AUC 0.50. The last considered study by Akari et al. [48] (R9) uses 56 features for 4004 patients and only applied the SVM model. The protocol used is K10. The results show an accuracy of 98.43 and a reliability index of 97.32%.

Our system AtheroEdge™ 3.0_{HDL} uses HDL systems that include the coronary angiography (gold standard) and ultrasound (contrast-enhanced, carotid B-mode) as part of the risk variables. Our novel model consists of three ML models, three unidirectional DL models, three bidirectional DL models, and 12 DL-based hybrid models. The results showed that the bidirectional DL models are better performing when compared to the unidirectional DL. There was around a **30.20%** improvement in the HDL models over the previous ML-based systems for stratification of CVD risk in the same Canadian cohort. We have also performed statistical tests to prove the reliability of the HDL systems. Furthermore, the top 10 features were selected by using three methods. Lastly, the validation was performed with the unseen dataset and obtained a ~1% difference in the results for the HDL systems.

8.2. A Special Note on Hybrid Deep Learning

Hybrid models increase overall performance in capturing complicated patterns by utilizing the advantages of many neural network designs. HDL models are adaptable to certain tasks, enabling the best elements to be chosen for each aspect of the issue. Transfer learning is made easier via the integration of pre-trained models, which allows for effective knowledge transfer between tasks and domains. HDL models function as ensembles, using several modelling techniques to minimize over fitting and improve resilience. They could provide improved interpretability since transparently constructed components make it easier to comprehend the judgments made via the model. HDL models may balance computing demands more effectively, which makes them appropriate for use in contexts with limited resources. HDL models use specific structures for each modality to manage multimodal data efficiently. Improvements in neural network research may be readily incorporated into HDL models, guaranteeing flexibility in response to new methods. Because of HDL models' flexibility, task-specific architectures may be designed to maximize performance for certain goals. HDL designs, which increase model applicability, frequently result in higher generalization on a variety of datasets by merging many models.

8.3. Strengths, Weaknesses, and Extensions

The HDL-based system was initially put forth. Focused carotid B-mode ultrasound, contrast-enhanced ultrasonography, and coronary angiography were all used in the AtheroEdge 3.0_{HDL} planned system. The HDL-based models have deep layers that produce a generalized training model when trained with epochs, optimal batch sizes, and learning rates. When compared to ML outcomes, this yields better HDL results. Fusing LBBM and OBBM with radionics-based features provided by MedUSE, CUSIP, and IPN features is another benefit of AtheroEdge 3.0_{HDL}. Three techniques were employed to choose the ML features to improve performance. The effect of different CV protocols was also studied for four types of AI models. The validation of the proposed HDL model was accomplished by applying it

to unseen dataset which showed a difference of 1%. The Canadian ethnicity datasets were considered for both seen and unseen data.

The following are the system's weaknesses and how they were mitigated: (i) *retraining of models*: the models need to be retrained in light of the altered input data. This modification might involve changing the sample size (the total number of patients) or adding new risk factors. There are ways to mitigate the effects even while retraining from scratch is necessary, such as by using transfer learning models; (ii) *sizes of training models*: AI models are always susceptible to having big training sets. Numerous factors are to blame for this. The AI techniques are under additional stress in terms of model sizes of the training patches due to the growth of risk factors and patient cohorts. Pruning AI algorithms is one mitigating strategy, notwithstanding the possibility of larger model sizes [101]; (iii) *AI bias*: due to a variety of factors, including bias in the prediction system, algorithm, and input datasets, bias in AI is a constant problem. Although bias can be present at many stages in the design of an AI system, there are mitigation techniques that allow us to identify and get rid of AI bias, including algorithms like Butterfly, PROBAST, and ROBINS methods [102]; (iv) *restrictions with hardware*: AI techniques that operate on the CPU clusters often require additional processing time. This depends on the quantity of data, batch sizes, learning rate, and epochs. GPU clusters allow for the mitigation of these variables, even if they have an impact on performance time [103]; (v) the proposed model applies only the CCE loss function [104–106]. Other loss function types can also be implemented. (vi) More performance comparisons can be conducted once the data sharing protocols are adopted.

In terms of improvements to the current study, distinct forms of plaque, such as symptomatic and asymptomatic plaque identified using the modality of grayscale type, could be included as one of the risks variable for CVD/stroke risk assessment [107–109]. Conventional approaches for wall segmentation can be used for CUSIP measurements [110–112]. In non-conventional approaches, sophisticated DL algorithms for CUSIP measurements can enhance the wall segmentation approaches [113–115]. The big data idea may be used in subsequent research to improve the accuracy of CVD risk forecasts [116–120]. To lower down the training sizes of the AI models, evolutionary approaches can also be used [47,101,121–123]. Some advanced models like (a) attention-based and (b) transformer-based models can be added to the underlying AI models [124–126], such as unidirectional and bidirectional models. While the above AI techniques utilize 2D ultrasound images, we can also extend this to 3D ultrasound [127–129]. *Additionally, we would like to attempt the combination of three models such as LSTM, RNN, and GRU via the nested fusion approach.*

9. Conclusions

Our technique is unique in that it employs HDL models with targeted carotid ultrasound (B-mode) as one of the risk predictors and contrast-enhanced ultrasonography with angiography of the coronary as the ground truth. Three machine learning models as well as three unidirectional, three bidirectional, and twelve hybrid models made up our system. Bidirectional models outperform unidirectional ones. Furthermore, using the same Canadian population, the HDL systems demonstrated a **30.20%** improvement over the prior ML-based method for CVD risk classification. Additionally, we have run the statistical tests necessary to demonstrate the HDL systems' dependability. Additionally, three methodologies were used to choose the top ten attributes. Finally, the validation was carried out using an unknown dataset, and the findings for the HDL systems showed a difference of less than 1%. The effect of data size was also studied showing the increase in the accuracy with the increasing K protocol. Further, different types of loss functions can be implemented in future work. The AtheroEdge 3.0_{HDL} system works in both an online and offline manner.

Author Contributions: Conceptualization, M.B. and J.S.S.; Methodology, M.B.; Software, M.B. and S.G.; Validation, S.P., L.M., A.M.J., J.R.L., I.M.S., N.N.K., M.A.-M., E.R.I., R.S., A.N., L.S., V.A. and J.S.S.; Formal analysis, S.G., N.N.K. and A.N.; Investigation, L.M., A.M.J., J.R.L., I.M.S., N.N.K., M.A.-M., E.R.I., R.S., A.N., L.S., V.A. and J.S.S.; Resources, S.P. and L.S.; Data curation, M.B.; Writing—original draft, M.B.; Writing—review & editing, M.B., S.G., E.T. and J.S.S.; Visualization, S.P., R.S., V.A. and J.S.S.; Supervision, S.P. and J.S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: There are no Institutional Review Board issues applicable to this study.

Informed Consent Statement: Not applicable in this study.

Data Availability Statement: Data are not available due to the proprietary nature of this study.

Acknowledgments: The authors would like to acknowledge AtheroPoint™, Roseville, CA 95661, USA for the medical data and the codes that are proprietary to AtheroPoint™.

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

Appendix A. Unidirectional Architecture

Table A1. Baseline Characteristics.

SN	Risk Factors or Target Variable	Samples (N) or Mean	Pr (%)
1	Age	64.49 years	-
2	Gender	349	69.8
3	Obesity	215	43.0
4	Ethnicity	486	97.2
5	BMI	31.12 kg/m ²	-
6	Hypertension	338	67.6
7	Angina	124	24.8
8	Diastolic blood pressure	76.7 mmHg	-
9	Systolic blood pressure	135.35 mmHg	-
10	Smoking history	330	66.0
11	Casual smoker	15	3.0
12	Current smoker	100	20.0
13	Previous smoker	218	43.6
14	Drinks	4.94 ± 10.4 per week	-
15	Family history of diabetics	195	39.0
16	Premature CVD in the family	146	29.2
17	CVD in the family	321	64.2
18	Pre-diabetic	20	40.0
19	Hyperlipidemia	288	57.6
20	Type II diabetes	114	22.8
21	Type I diabetes	5	1.0
22	Creatinine	83.99 ± 22.6 µmol/L	-
23	Estimated glomerular filtration rate	78.96 mL/min/1.73 m ²	-
24	Diabetes of any type	118	23.6
25	Total plaque area (TPA)	47.68 mm ²	-

Table A1. Cont.

SN	Risk Factors or Target Variable	Samples (N) or Mean	Pr (%)
26	Maximum plaque height (MPH)	2.64 mm	-
27	Intra-plaque neovascularization (IPN)	1.16	-
28	Angiotensin-converting enzyme (ACE) inhibitors	191	38.2
29	HMG-Co reductase inhibitors	272	54.4
30	Angiotensin receptor blockers (ARBs)	45	9.0
31	Other antilipemic agents	9	1.80
32	Alpha-blockers	30	6.0
33	Calcium channel-blockers	93	18.6
34	Beta-blockers	236	47.2
35	Diuretics	99	19.8
36	Anti-platelet	368	73.6
37	Insulin	38	7.6
38	Anti-anginals and NSAIDS	81	16.20
39	Non-insulin diabetes medications	72	14.4

Appendix B. Unidirectional Architecture

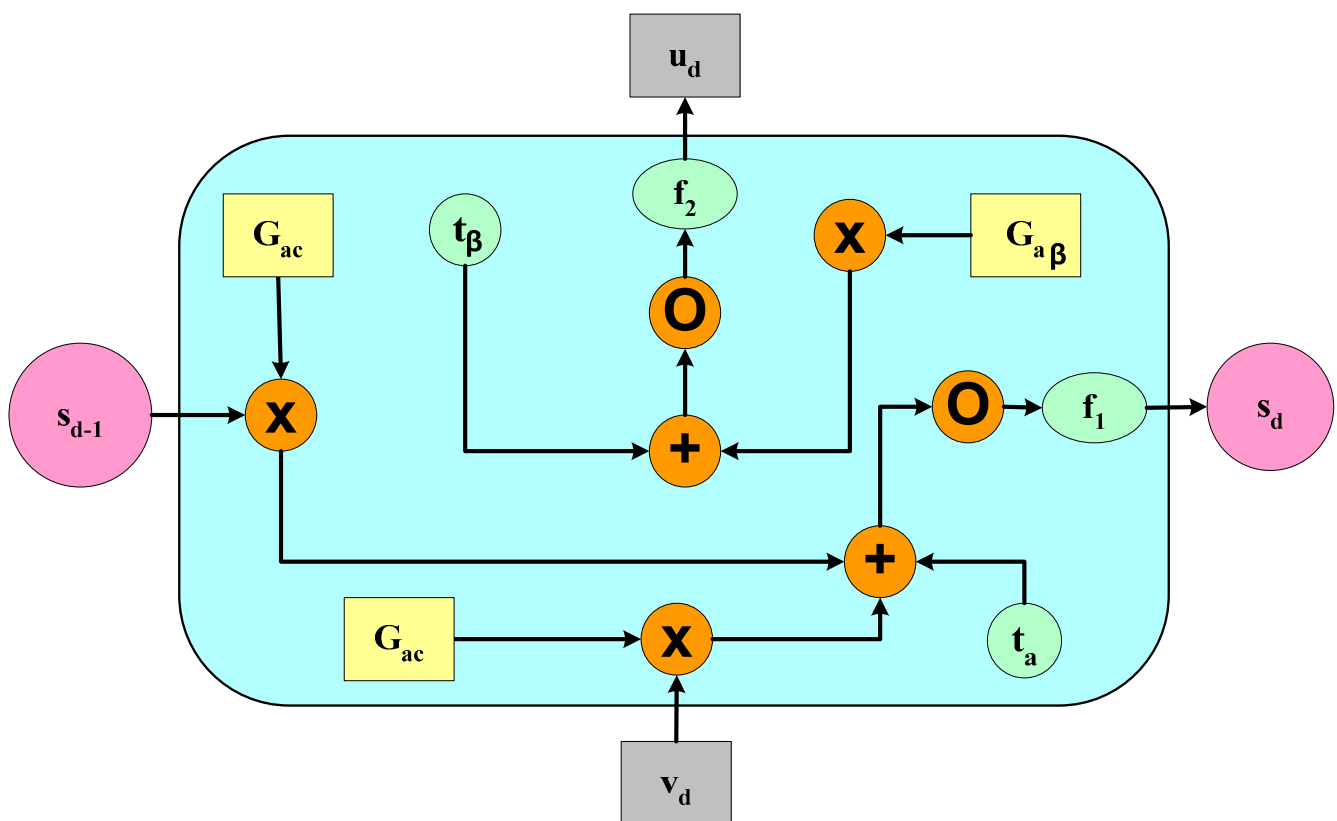


Figure A1. RNN Architecture.

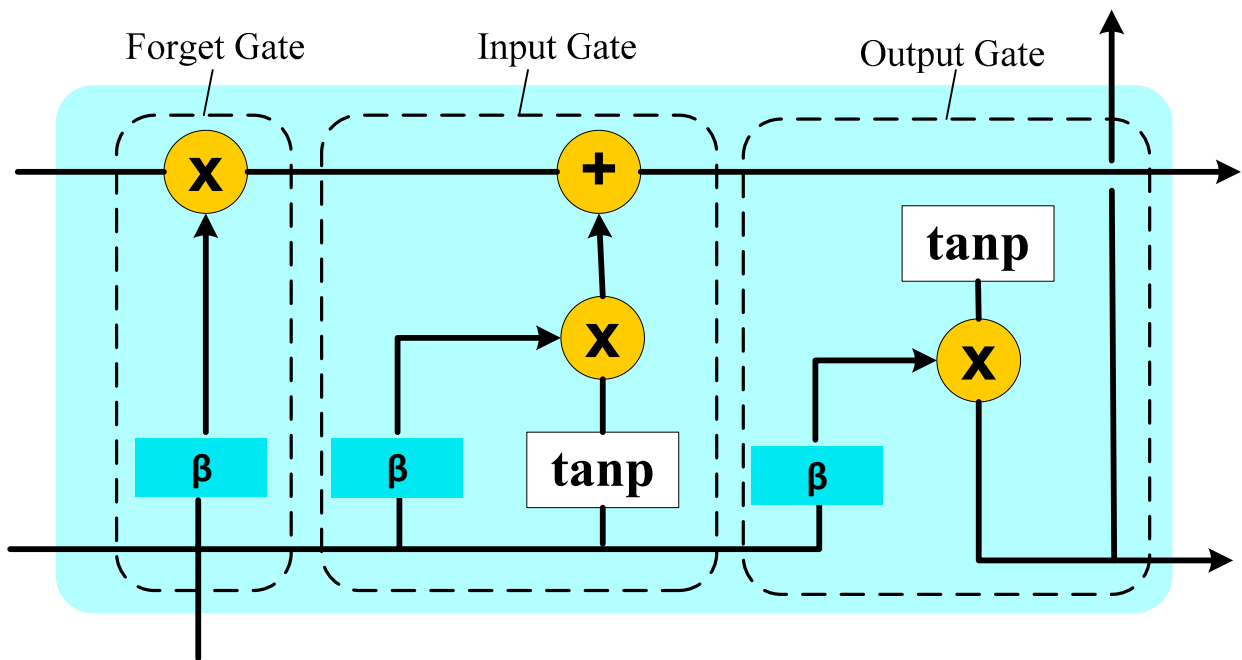


Figure A2. LSTM Architecture.

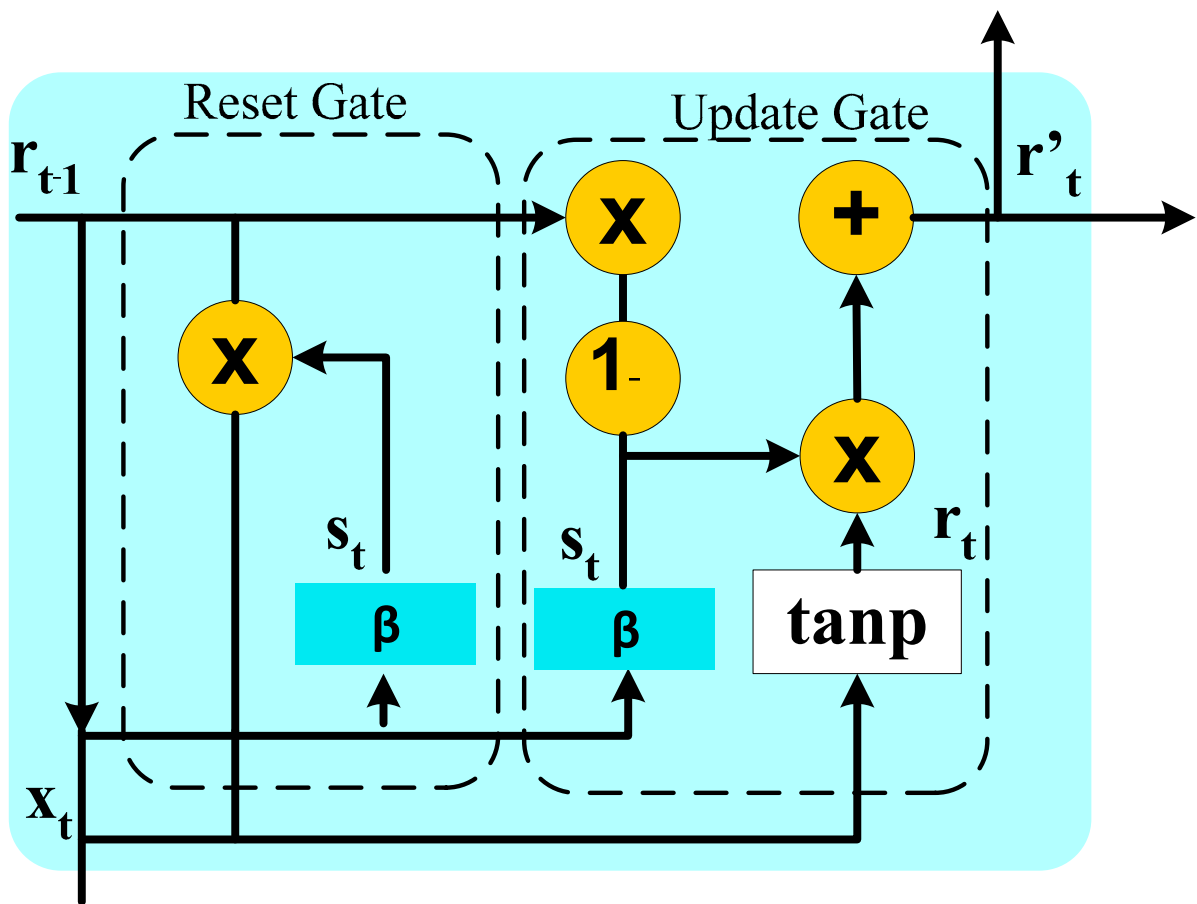


Figure A3. GRU Architecture.

Appendix C. Bidirectional Architecture

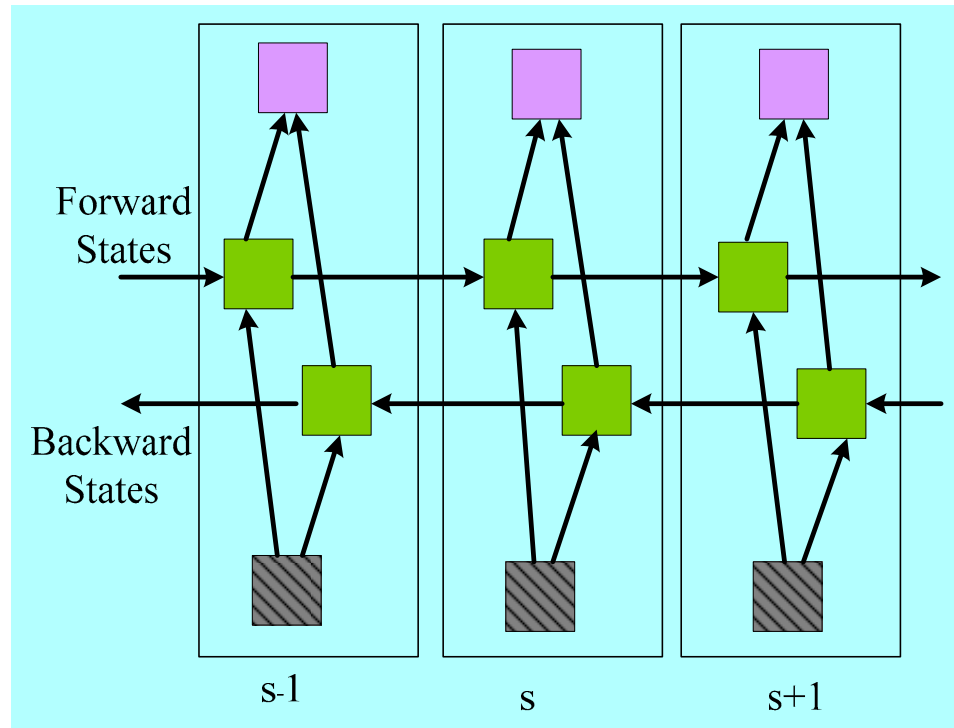


Figure A4. BiRNN Architecture.

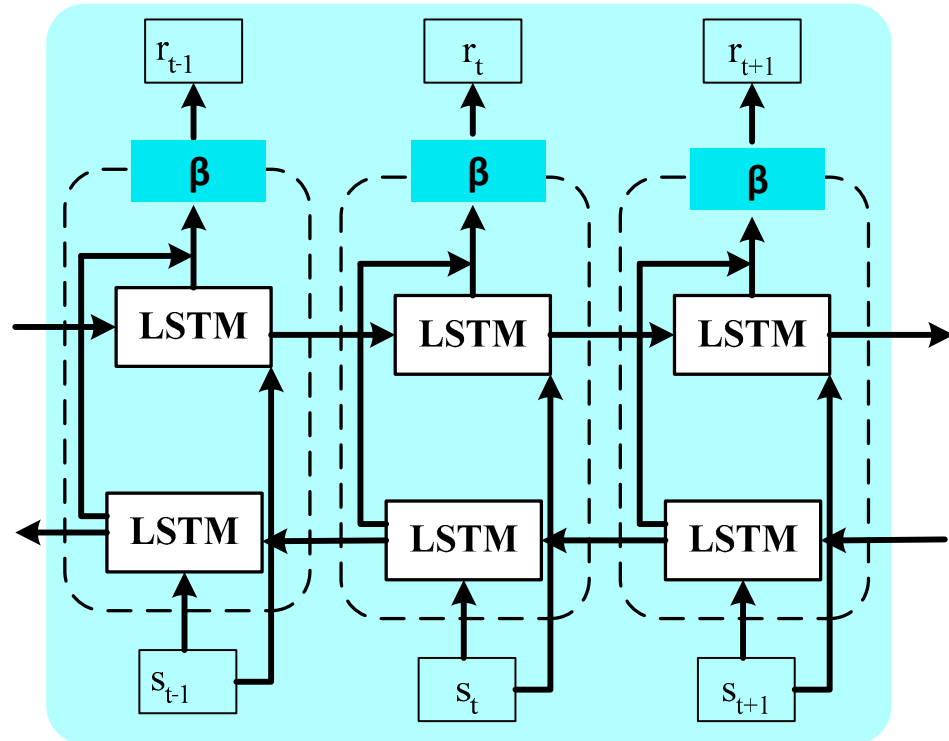


Figure A5. BiLSTM Architecture.

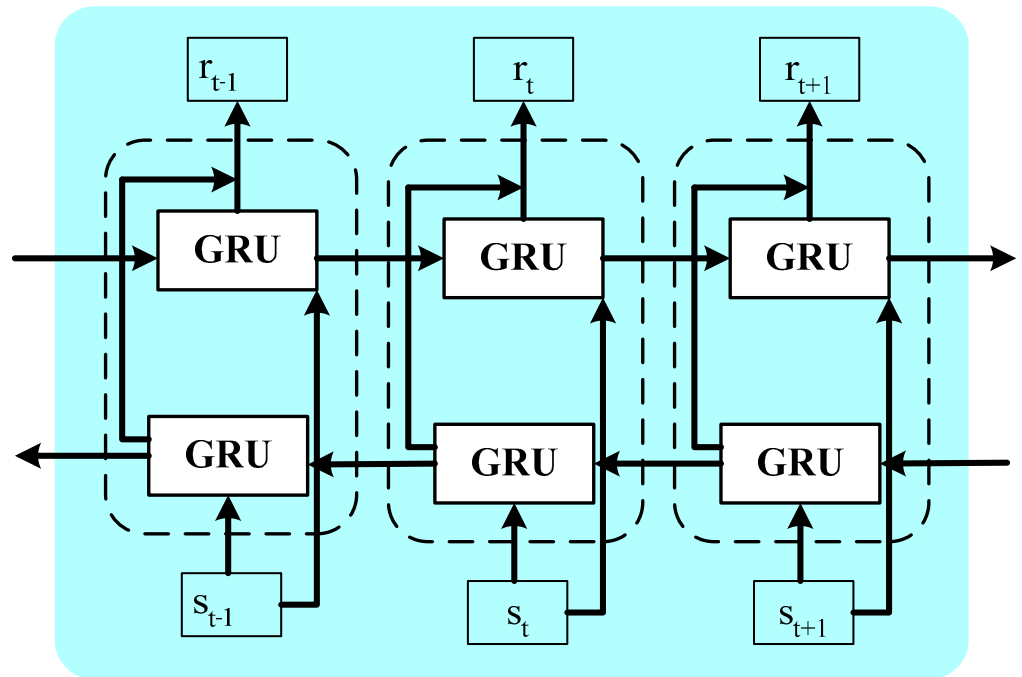


Figure A6. BiGRU Architecture.

Appendix D. Hybrid Deep Learning Architecture

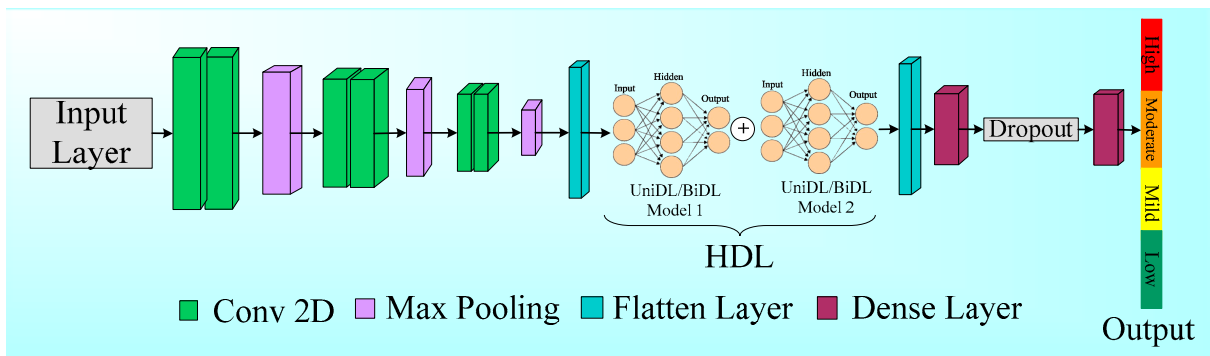


Figure A7. Hybrid deep learning architecture; UniDL: unidirectional DL; BiDL: bidirectional DL.

Appendix E. Scientific Validation with the Unseen Data

Table A2. Scientific validation with the unseen data.

SN	Unseen Data (Train A-Test B)			Unseen Data (Train B-Test A)	
	Models	Mean AUC	<i>p</i> -Value	Mean AUC	<i>p</i> -Value
1	ML	0.683	<0.005	0.683	<0.005
2	UniDL	0.884	<0.001	0.880	<0.001
3	BiDL	0.905	<0.001	0.900	<0.001
4	HDL	0.947	<0.001	0.940	<0.001

SN: Serial Number; ML: Machine Learning; DL: Deep Learning; UniDL: Unidirectional DL; BiDL: Bidirectional DL; AUC: Area-Under-the-Curve; HDL: Hybrid DL.

Appendix F. Confusion Matrix for Unidirectional HDL and Bidirectional HDL Model

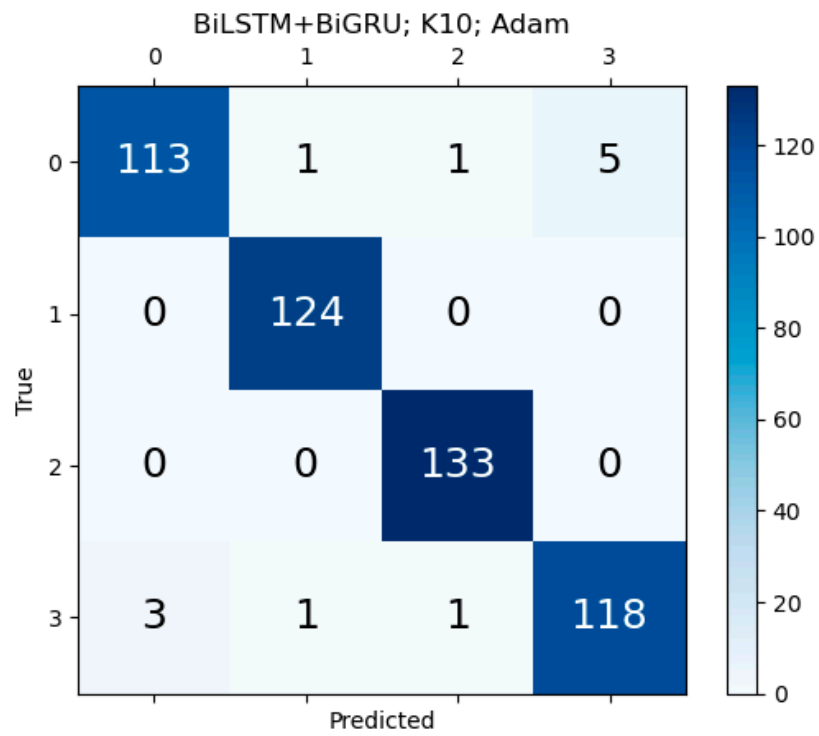


Figure A8. Confusion matrix for the best K10 protocol which fuses BiLSTM and BiGRU.

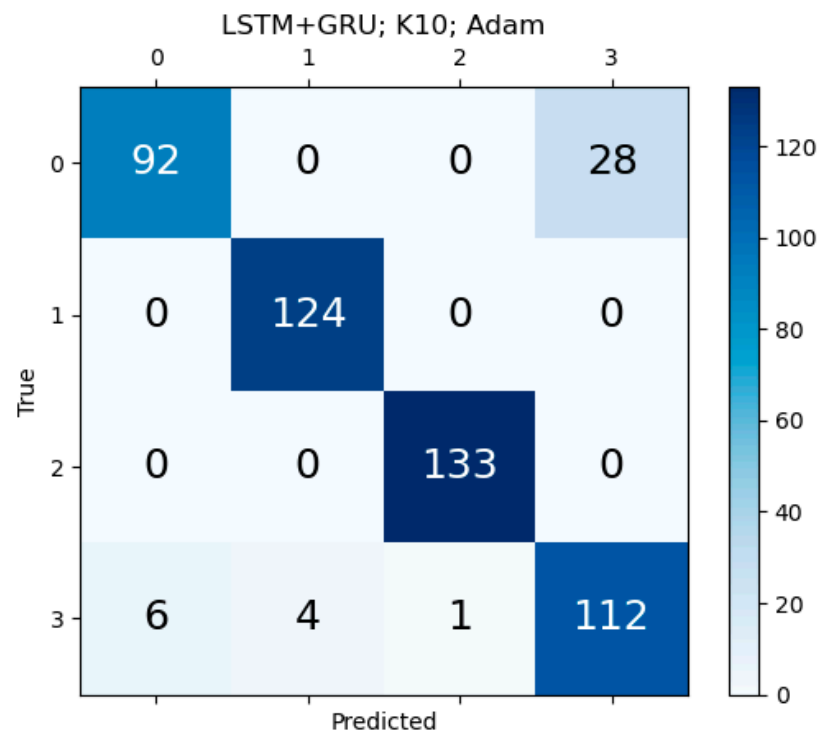


Figure A9. Confusion matrix for the best K10 protocol which fuses LSTM and GRU.

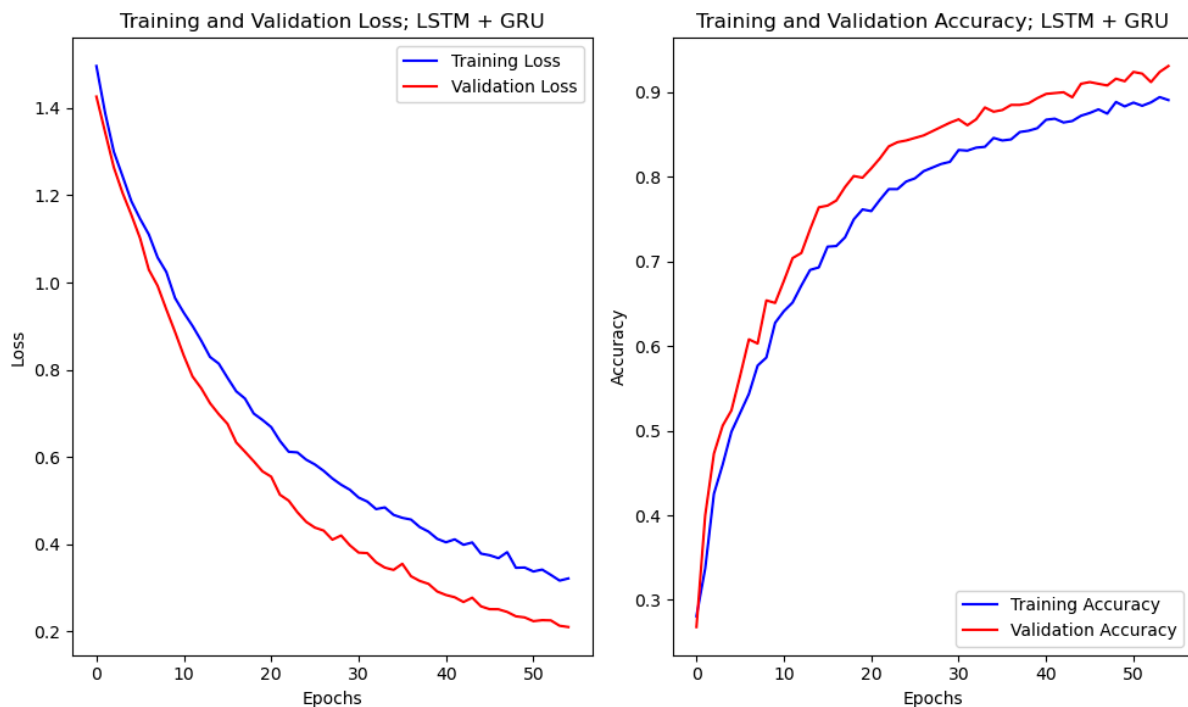


Figure A10. Loss and accuracy curves for the best K10 protocol which fuses LSTM and GRU.

References

1. Kaptoge, S.; Pennells, L.; De Bacquer, D.; Cooney, M.T.; Kavousi, M.; Stevens, G.; Riley, L.M.; Savin, S.; Khan, T.; Altay, S.; et al. World Health Organization cardiovascular disease risk charts: Revised models to estimate risk in 21 global regions. *Lancet Glob. Health* **2019**, *7*, e1332–e1345. [[CrossRef](#)] [[PubMed](#)]
2. Suri, J.S.; Agarwal, S.; Gupta, S.K.; Puvvula, A.; Biswas, M.; Saba, L.; Bit, A.; Tandel, G.S.; Agarwal, M.; Patrick, A.; et al. A narrative review on characterization of acute respiratory distress syndrome in COVID-19-infected lungs using artificial intelligence. *Comput. Biol. Med.* **2021**, *130*, 104210. [[CrossRef](#)] [[PubMed](#)]
3. Jamthikar, A.D.; Puvvula, A.; Gupta, D.; Johri, A.M.; Nambi, V.; Khanna, N.N.; Saba, L.; Mavrogeni, S.; Laird, J.R.; Pareek, G. Cardiovascular disease and stroke risk assessment in patients with chronic kidney disease using integration of estimated glomerular filtration rate, ultrasonic image phenotypes, and artificial intelligence: A narrative review. *Int. Angiol. A J. Int. Union Angiol.* **2020**, *40*, 150–164. [[CrossRef](#)]
4. Saba, L.; Anzidei, M.; Sanfilippo, R.; Montisci, R.; Lucatelli, P.; Catalano, C.; Passariello, R.; Mallarini, G. Imaging of the carotid artery. *Atherosclerosis* **2012**, *220*, 294–309. [[CrossRef](#)]
5. Griffin, M.; Nicolaidis, A.N.; Belcaro, G.; Shah, E. Cardiovascular risk assessment using ultrasound: The value of arterial wall changes including the presence, severity and character of plaques. *Pathophysiol. Haemost. Thromb.* **2002**, *32*, 367–370. [[CrossRef](#)]
6. Suri, J.S.; Kathuria, C.; Molinari, F. *Atherosclerosis Disease Management*; Springer: Berlin/Heidelberg, Germany, 2010.
7. Saba, L.; Jamthikar, A.; Gupta, D.; Khanna, N.N.; Viskovic, K.; Suri, H.S.; Gupta, A.; Mavrogeni, S.; Turk, M.; Laird, J.R. Global perspective on carotid intima-media thickness and plaque: Should the current measurement guidelines be revisited? *Int. Angiol.* **2019**, *38*, 451–465. [[CrossRef](#)]
8. Giannopoulos, A.A.; Kyriacou, E.; Griffin, M.; Pattichis, C.S.; Michael, J.; Richards, T.; Geroulakos, G.; Nicolaidis, A.N. Dynamic carotid plaque imaging using ultrasonography. *J. Vasc. Surg.* **2021**, *73*, 1630–1638. [[CrossRef](#)]
9. Amato, M.; Montorsi, P.; Ravani, A.; Oldani, E.; Galli, S.; Ravagnani, P.M.; Tremoli, E.; Baldassarre, D. Carotid intima-media thickness by B-mode ultrasound as surrogate of coronary atherosclerosis: Correlation with quantitative coronary angiography and coronary intravascular ultrasound findings. *Eur. Heart J.* **2007**, *28*, 2094–2101. [[CrossRef](#)]
10. Bots, M.L. Carotid intima-media thickness as a surrogate marker for cardiovascular disease in intervention studies. *Curr. Med. Res. Opin.* **2006**, *22*, 2181–2190. [[CrossRef](#)]
11. Spence, J.D. Ultrasound measurement of carotid plaque as a surrogate outcome for coronary artery disease. *Am. J. Cardiol.* **2002**, *89*, 10–15. [[CrossRef](#)]
12. Puvvula, A.; Jamthikar, A.D.; Gupta, D.; Khanna, N.N.; Porcu, M.; Saba, L.; Viskovic, K.; Ajuluchukwu, J.N.; Gupta, A.; Mavrogeni, S. Morphological carotid plaque area is associated with glomerular filtration rate: A study of south asian indian patients with diabetes and chronic kidney disease. *Angiology* **2020**, *71*, 520–535. [[CrossRef](#)]
13. Biswas, M.; Saba, L.; Chakrabartty, S.; Khanna, N.N.; Song, H.; Suri, H.S.; Sfikakis, P.P.; Mavrogeni, S.; Viskovic, K.; Laird, J.R. Two-stage artificial intelligence model for jointly measurement of atherosclerotic wall thickness and plaque burden in carotid ultrasound: A screening tool for cardiovascular/stroke risk assessment. *Comput. Biol. Med.* **2020**, *123*, 103847. [[CrossRef](#)]

14. Landry, A.; Spence, J.D.; Fenster, A. Measurement of carotid plaque volume by 3-dimensional ultrasound. *Stroke* **2004**, *35*, 864–869. [[CrossRef](#)]
15. Johri, A.M.; Lajkosz, K.A.; Grubic, N.; Islam, S.; Li, T.Y.; Simpson, C.S.; Ewart, P.; Suri, J.S.; Héту, M.-F. Maximum plaque height in carotid ultrasound predicts cardiovascular disease outcomes: A population-based validation study of the American society of echocardiography's grade II–III plaque characterization and protocol. *Int. J. Cardiovasc. Imaging* **2021**, *37*, 1601–1610. [[CrossRef](#)]
16. Mantella, L.E.; Colledanchise, K.N.; Hetu, M.-F.; Feinstein, S.B.; Abunassar, J.; Johri, A.M. Carotid intraplaque neovascularization predicts coronary artery disease and cardiovascular events. *Eur. Heart J. Cardiovasc. Imaging* **2019**, *20*, 1239–1247. [[CrossRef](#)]
17. Goldstein, B.A.; Navar, A.M.; Carter, R.E. Moving beyond regression techniques in cardiovascular risk prediction: Applying machine learning to address analytic challenges. *Eur. Heart J.* **2017**, *38*, 1805–1814. [[CrossRef](#)]
18. Jamthikar, A.D.; Gupta, D.; Saba, L.; Khanna, N.N.; Viskovic, K.; Mavrogeni, S.; Laird, J.R.; Sattar, N.; Johri, A.M.; Pareek, G. Artificial intelligence framework for predictive cardiovascular and stroke risk assessment models: A narrative review of integrated approaches using carotid ultrasound. *Comput. Biol. Med.* **2020**, *126*, 104043. [[CrossRef](#)]
19. Alaa, A.M.; Bolton, T.; Di Angelantonio, E.; Rudd, J.H.; Van der Schaar, M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS ONE* **2019**, *14*, e0213653. [[CrossRef](#)] [[PubMed](#)]
20. Weng, S.F.; Reys, J.; Kai, J.; Garibaldi, J.M.; Qureshi, N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS ONE* **2017**, *12*, e0174944. [[CrossRef](#)]
21. Biswas, M.; Kuppili, V.; Saba, L.; Edla, D.R.; Suri, H.S.; Cuadrado-Godia, E.; Laird, J.R.; Marinhe, R.T.; Sanches, J.M.; Nicolaidis, A. State-of-the-art review on deep learning in medical imaging. *Front. Biosci. Landmark* **2019**, *24*, 380–406.
22. Saba, L.; Biswas, M.; Kuppili, V.; Godia, E.C.; Suri, H.S.; Edla, D.R.; Omerzu, T.; Laird, J.R.; Khanna, N.N.; Mavrogeni, S. The present and future of deep learning in radiology. *Eur. J. Radiol.* **2019**, *114*, 14–24. [[CrossRef](#)] [[PubMed](#)]
23. Hosny, A.; Parmar, C.; Quackenbush, J.; Schwartz, L.H.; Aerts, H.J. Artificial intelligence in radiology. *Nat. Rev. Cancer* **2018**, *18*, 500–510. [[CrossRef](#)] [[PubMed](#)]
24. Pianykh, O.S.; Lings, G.; Dewey, M.; Enzmann, D.R.; Herold, C.J.; Schoenberg, S.O.; Brink, J.A. Continuous learning AI in radiology: Implementation principles and early applications. *Radiology* **2020**, *297*, 6–14. [[CrossRef](#)]
25. Thrall, J.H.; Li, X.; Li, Q.; Cruz, C.; Do, S.; Dreyer, K.; Brink, J. Artificial intelligence and machine learning in radiology: Opportunities, challenges, pitfalls, and criteria for success. *J. Am. Coll. Radiol.* **2018**, *15*, 504–508. [[CrossRef](#)] [[PubMed](#)]
26. Shrivastava, V.K.; Londhe, N.D.; Sonawane, R.S.; Suri, J.S. Reliable and accurate psoriasis disease classification in dermatology images using comprehensive feature space in machine learning paradigm. *Expert Syst. Appl.* **2015**, *42*, 6184–6195. [[CrossRef](#)]
27. Shrivastava, V.K.; Londhe, N.D.; Sonawane, R.S.; Suri, J.S. A novel and robust Bayesian approach for segmentation of psoriasis lesions and its risk stratification. *Comput. Methods Programs Biomed.* **2017**, *150*, 9–22. [[CrossRef](#)]
28. Du-Harpur, X.; Watt, F.; Luscombe, N.; Lynch, M. What is AI? Applications of artificial intelligence to dermatology. *Br. J. Dermatol.* **2020**, *183*, 423–430. [[CrossRef](#)]
29. Li, C.-X.; Shen, C.-B.; Xue, K.; Shen, X.; Jing, Y.; Wang, Z.-Y.; Xu, F.; Meng, R.-S.; Yu, J.-B.; Cui, Y. Artificial intelligence in dermatology: Past, present, and future. *Chin. Med. J.* **2019**, *132*, 2017–2020. [[CrossRef](#)]
30. Fritzsche, K.; Can, A.; Shen, H.; Tsai, C.; Turner, J.; Tanenbaum, H.; Stewart, C.; Roysam, B.; Suri, J.; Laxminarayan, S. Automated model based segmentation, tracing and analysis of retinal vasculature from digital fundus images. In *State-of-The-Art Angiography, Applications and Plaque Imaging Using MR, CT, Ultrasound and X-rays*; CRC Press: Boca Raton, FL, USA, 2003; Volume 29, pp. 225–298.
31. Hogarty, D.T.; Mackey, D.A.; Hewitt, A.W. Current state and future prospects of artificial intelligence in ophthalmology: A review. *Clin. Exp. Ophthalmol.* **2019**, *47*, 128–139. [[CrossRef](#)]
32. Tong, Y.; Yu, Y.; Xing, Y.; Chen, C.; Shen, Y. Applications of artificial intelligence in ophthalmology: General overview. *J. Ophthalmol.* **2018**, *2018*, 5278196.
33. Ting, D.S.W.; Pasquale, L.R.; Peng, L.; Campbell, J.P.; Lee, A.Y.; Raman, R.; Tan, G.S.W.; Schmetterer, L.; Keane, P.A.; Wong, T.Y. Artificial intelligence and deep learning in ophthalmology. *Br. J. Ophthalmol.* **2019**, *103*, 167–175. [[CrossRef](#)] [[PubMed](#)]
34. Schmidt-Erfurth, U.; Reiter, G.S.; Riedl, S.; Seeböck, P.; Vogl, W.-D.; Blodi, B.A.; Domalpally, A.; Fawzi, A.; Jia, Y.; Sarraf, D. AI-based monitoring of retinal fluid in disease activity and under therapy. *Prog. Retin. Eye Res.* **2022**, *86*, 100972. [[CrossRef](#)]
35. Sorrentino, F.S.; Jurman, G.; De Nadai, K.; Campa, C.; Furlanello, C.; Parmeggiani, F. Application of artificial intelligence in targeting retinal diseases. *Curr. Drug Targets* **2020**, *21*, 1208–1215. [[CrossRef](#)] [[PubMed](#)]
36. Saba, L.; Sanfilippo, R.; Sanna, S.; Anzidei, M.; Montisci, R.; Mallarini, G.; Suri, J.S. Association between carotid artery plaque volume, composition, and ulceration: A retrospective assessment with MDCT. *Am. J. Roentgenol.* **2012**, *199*, 151–156. [[CrossRef](#)]
37. Acharya, U.R.; Sree, S.V.; Krishnan, M.M.R.; Krishnananda, N.; Ranjan, S.; Umesh, P.; Suri, J.S. Automated classification of patients with coronary artery disease using grayscale features from left ventricle echocardiographic images. *Comput. Methods Programs Biomed.* **2013**, *112*, 624–632. [[CrossRef](#)]
38. Johnson, K.W.; Soto, J.T.; Glicksberg, B.S.; Shameer, K.; Miotto, R.; Ali, M.; Ashley, E.; Dudley, J.T. Artificial intelligence in cardiology. *J. Am. Coll. Cardiol.* **2018**, *71*, 2668–2679. [[CrossRef](#)]
39. Lopez-Jimenez, F.; Attia, Z.; Arruda-Olson, A.M.; Carter, R.; Chareonthaitawee, P.; Jouni, H.; Kapa, S.; Lerman, A.; Luong, C.; Medina-Inojosa, J.R. Artificial intelligence in cardiology: Present and future. In *Mayo Clinic Proceedings*; Elsevier: Amsterdam, The Netherlands, 2020.

40. Molinari, F.; Mantovani, A.; Deandrea, M.; Limone, P.; Garberoglio, R.; Suri, J.S. Characterization of single thyroid nodules by contrast-enhanced 3-D ultrasound. *Ultrasound Med. Biol.* **2010**, *36*, 1616–1625. [[CrossRef](#)]
41. Gubbi, S.; Hamet, P.; Tremblay, J.; Koch, C.A.; Hannah-Shmouni, F. Artificial intelligence and machine learning in endocrinology and metabolism: The dawn of a new era. *Front. Endocrinol.* **2019**, *10*, 185. [[CrossRef](#)]
42. Giorgini, F.; Di Dalmazi, G.; Diciotti, S. Artificial intelligence in endocrinology: A comprehensive review. *J. Endocrinol. Investig.* **2024**, *47*, 1067–1082. [[CrossRef](#)]
43. Thomasian, N.M.; Kamel, I.R.; Bai, H.X. Machine intelligence in non-invasive endocrine cancer diagnostics. *Nat. Rev. Endocrinol.* **2022**, *18*, 81–95. [[CrossRef](#)]
44. Jamthikar, A.; Gupta, D.; Khanna, N.N.; Saba, L.; Araki, T.; Viskovic, K.; Suri, H.S.; Gupta, A.; Mavrogeni, S.; Turk, M. A low-cost machine learning-based cardiovascular/stroke risk assessment system: Integration of conventional factors with image phenotypes. *Cardiovasc. Diagn. Ther.* **2019**, *9*, 420. [[CrossRef](#)]
45. Kakadiaris, I.A.; Vrigkas, M.; Yen, A.A.; Kuznetsova, T.; Budoff, M.; Naghavi, M. Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *J. Am. Heart Assoc.* **2018**, *7*, e009476. [[CrossRef](#)]
46. Arsenescu, T.; Chifor, R.; Marita, T.; Santoma, A.; Lebovici, A.; Duma, D.; Vacaras, V.; Badea, A.F. 3D ultrasound reconstructions of the carotid artery and thyroid gland using artificial-intelligence-based automatic segmentation—Qualitative and quantitative evaluation of the segmentation results via comparison with CT angiography. *Sensors* **2023**, *23*, 2806. [[CrossRef](#)] [[PubMed](#)]
47. Acharya, U.R.; Mookiah, M.R.K.; Sree, S.V.; Yanti, R.; Martis, R.; Saba, L.; Molinari, F.; Guerriero, S.; Suri, J.S. Evolutionary algorithm-based classifier parameter tuning for automatic ovarian cancer tissue characterization and classification. *Ultraschall Med. Eur. J. Ultrasound* **2014**, *35*, 237–245.
48. Araki, T.; Ikeda, N.; Shukla, D.; Jain, P.K.; Londhe, N.D.; Shrivastava, V.K.; Banchhor, S.K.; Saba, L.; Nicolaidis, A.; Shafique, S. PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: A link between carotid and coronary grayscale plaque morphology. *Comput. Methods Programs Biomed.* **2016**, *128*, 137–158. [[CrossRef](#)] [[PubMed](#)]
49. Anaya-Isaza, A.; Mera-Jiménez, L.; Zequera-Diaz, M. An overview of deep learning in medical imaging. *Inform. Med. Unlocked* **2021**, *26*, 100723. [[CrossRef](#)]
50. Lee, J.-G.; Jun, S.; Cho, Y.-W.; Lee, H.; Kim, G.B.; Seo, J.B.; Kim, N. Deep learning in medical imaging: General overview. *Korean J. Radiol.* **2017**, *18*, 570–584. [[CrossRef](#)]
51. Aggarwal, R.; Sounderajah, V.; Martin, G.; Ting, D.S.; Karthikesalingam, A.; King, D.; Ashrafian, H.; Darzi, A. Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis. *NPJ Digit. Med.* **2021**, *4*, 65. [[CrossRef](#)]
52. Khanna, N.N.; Maindarkar, M.A.; Viswanathan, V.; Puvvula, A.; Paul, S.; Bhagawati, M.; Ahluwalia, P.; Ruzsa, Z.; Sharma, A.; Kolluri, R. Cardiovascular/Stroke Risk Stratification in Diabetic Foot Infection Patients Using Deep Learning-Based Artificial Intelligence: An Investigative Study. *J. Clin. Med.* **2022**, *11*, 6844. [[CrossRef](#)]
53. Suri, J.S.; Bhagawati, M.; Paul, S.; Protogerou, A.D.; Sfikakis, P.P.; Kitas, G.D.; Khanna, N.N.; Ruzsa, Z.; Sharma, A.M.; Saxena, S. A powerful paradigm for cardiovascular risk stratification using multiclass, multi-label, and ensemble-based machine learning paradigms: A narrative review. *Diagnostics* **2022**, *12*, 722. [[CrossRef](#)] [[PubMed](#)]
54. Johri, A.M.; Singh, K.V.; Mantella, L.E.; Saba, L.; Sharma, A.; Laird, J.R.; Utkarsh, K.; Singh, I.M.; Gupta, S.; Kalra, M.S. Deep learning artificial intelligence framework for multiclass coronary artery disease prediction using combination of conventional risk factors, carotid ultrasound, and intraplaque neovascularization. *Comput. Biol. Med.* **2022**, *150*, 106018. [[CrossRef](#)] [[PubMed](#)]
55. Konstantonis, G.; Singh, K.V.; Sfikakis, P.P.; Jamthikar, A.D.; Kitas, G.D.; Gupta, S.K.; Saba, L.; Verrou, K.; Khanna, N.N.; Ruzsa, Z. Cardiovascular disease detection using machine learning and carotid/femoral arterial imaging frameworks in rheumatoid arthritis patients. *Rheumatol. Int.* **2022**, *42*, 215–239. [[CrossRef](#)]
56. Jamthikar, A.D.; Gupta, D.; Mantella, L.E.; Saba, L.; Laird, J.R.; Johri, A.M.; Suri, J.S. Multiclass machine learning vs. conventional calculators for stroke/CVD risk assessment using carotid plaque predictors with coronary angiography scores as gold standard: A 500 participants study. *Int. J. Cardiovasc. Imaging* **2021**, *37*, 1171–1187. [[CrossRef](#)] [[PubMed](#)]
57. Johri, A.M.; Mantella, L.E.; Jamthikar, A.D.; Saba, L.; Laird, J.R.; Suri, J.S. Role of artificial intelligence in cardiovascular risk prediction and outcomes: Comparison of machine-learning and conventional statistical approaches for the analysis of carotid ultrasound features and intra-plaque neovascularization. *Int. J. Cardiovasc. Imaging* **2021**, *37*, 3145–3156. [[CrossRef](#)]
58. Wilson, P.W.; D’Agostino, R.B.; Levy, D.; Belanger, A.M.; Silbershatz, H.; Kannel, W.B. Prediction of coronary heart disease using risk factor categories. *Circulation* **1998**, *97*, 1837–1847. [[CrossRef](#)]
59. Conroy, R.M.; Pyörälä, K.; Fitzgerald, A.E.; Sans, S.; Menotti, A.; De Backer, G.; De Bacquer, D.; Ducimetiere, P.; Jousilahti, P.; Keil, U. Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. *Eur. Heart J.* **2003**, *24*, 987–1003. [[CrossRef](#)]
60. Golf, D.C., Jr.; Lloyd-Jones, D.M.; Bennett, G.; Coady, S.; D’agostino, R.B.; Gibbons, R.; Greenland, P.; Lackland, D.T.; Levy, D.; O’donnell, C.J. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* **2014**, *129* (Suppl. S2), S49–S73.
61. Kumar, K.; Araki, T.; Rajan, J.; Saba, L.; Lavra, F.; Ikeda, N.; Sharma, A.M.; Shafique, S.; Nicolaidis, A.; Laird, J.R. Accurate lumen diameter measurement in curved vessels in carotid ultrasound: An iterative scale-space and spatial transformation approach. *Med. Biol. Eng. Comput.* **2017**, *55*, 1415–1434. [[CrossRef](#)]

62. Detrano, R.; Guerci, A.D.; Carr, J.J.; Bild, D.E.; Burke, G.; Folsom, A.R.; Liu, K.; Shea, S.; Szklo, M.; Bluemke, D.A. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. *N. Engl. J. Med.* **2008**, *358*, 1336–1345. [[CrossRef](#)]
63. Poplin, R.; Varadarajan, A.V.; Blumer, K.; Liu, Y.; McConnell, M.V.; Corrado, G.S.; Peng, L.; Webster, D.R. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat. Biomed. Eng.* **2018**, *2*, 158–164. [[CrossRef](#)]
64. Escaned, J.; Baptista, J.; Di Mario, C.; Haase, J.R.; Ozaki, Y.; Linker, D.T.; De Feyter, P.J.; Roelandt, J.R.; Serruys, P.W. Significance of automated stenosis detection during quantitative angiography: Insights gained from intracoronary ultrasound imaging. *Circulation* **1996**, *94*, 966–972. [[CrossRef](#)] [[PubMed](#)]
65. Bourantas, C.V.; Kalatzis, F.G.; Papafaklis, M.I.; Fotiadis, D.I.; Tweddel, A.C.; Kourtis, I.C.; Katsouras, C.S.; Michalis, L.K. ANGIOCARE: An automated system for fast three-dimensional coronary reconstruction by integrating angiographic and intracoronary ultrasound data. *Catheter. Cardiovasc. Interv.* **2008**, *72*, 166–175. [[CrossRef](#)]
66. Joseph, J.; Kiran, R.; Nabeel, P.; Shah, M.I.; Bhaskar, A.; Ganesh, C.; Seshadri, S.; Sivaprakasam, M. ARTSENS® Pen—Portable easy-to-use device for carotid stiffness measurement: Technology validation and clinical-utility assessment. *Biomed. Phys. Eng. Express* **2020**, *6*, 025013. [[CrossRef](#)]
67. Daigle, R.J. *Techniques in Noninvasive Vascular Diagnosis: An Encyclopedia of Vascular Testing*; Summer Publishing LLC: Littleton, CO, USA, 2008.
68. Nicolaides, A.N.; Kakkos, S.K.; Griffin, M.; Sabetai, M.; Dhanjil, S.; Thomas, D.J.; Geroulakos, G.; Georgiou, N.; Francis, S.; Ioannidou, E. Effect of image normalization on carotid plaque classification and the risk of ipsilateral hemispheric ischemic events: Results from the asymptomatic carotid stenosis and risk of stroke study. *Vascular* **2005**, *13*, 211–221. [[CrossRef](#)] [[PubMed](#)]
69. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]
70. Elreedy, D.; Atiya, A.F. A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance. *Inf. Sci.* **2019**, *505*, 32–64. [[CrossRef](#)]
71. Santoso, B.; Wijayanto, H.; Notodiputro, K.A.; Sartono, B. Synthetic over sampling methods for handling class imbalanced problems: A review. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2017.
72. Fernández, A.; Garcia, S.; Herrera, F.; Chawla, N.V. SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *J. Artif. Intell. Res.* **2018**, *61*, 863–905. [[CrossRef](#)]
73. Blagus, R.; Lusa, L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinform.* **2013**, *14*, 1–16. [[CrossRef](#)]
74. Deepa, B.; Ramesh, K. Epileptic seizure detection using deep learning through min max scaler normalization. *Int. J. Health Sci.* **2022**, *6*, 10981–10996. [[CrossRef](#)]
75. Sembiring, I.; Wahyuni, S.N.; Sedyono, E. LSTM algorithm optimization for COVID-19 prediction model. *Heliyon* **2024**, *10*, e26158. [[CrossRef](#)]
76. O'Donncha, F.; Hu, Y.; Palmes, P.; Burke, M.; Filgueira, R.; Grant, J. A spatio-temporal LSTM model to forecast across multiple temporal and spatial scales. *Ecol. Inform.* **2022**, *69*, 101687. [[CrossRef](#)]
77. Olhosseiny, H.H.; Mirzaloo, M.; Bolic, M.; Dajani, H.R.; Groza, V.; Yoshida, M. Identifying high risk of atherosclerosis using deep learning and ensemble learning. In Proceedings of the 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Lausanne, Switzerland, 23–25 June 2021.
78. An, Y.; Huang, N.; Chen, X.; Wu, F.; Wang, J. High-risk prediction of cardiovascular diseases via attention-based deep neural networks. *IEEE ACM Trans. Comput. Biol. Bioinform.* **2019**, *18*, 1093–1105. [[CrossRef](#)] [[PubMed](#)]
79. Baccouche, A.; Garcia-Zapirain, B.; Olea, C.C.; Elmaghraby, A. Ensemble deep learning models for heart disease classification: A case study from Mexico. *Information* **2020**, *11*, 207. [[CrossRef](#)]
80. Bai, S.; Yan, M.; Wan, Q.; He, L.; Wang, X.; Li, J. DL-RNN: An accurate indoor localization method via double RNNs. *IEEE Sens. J.* **2019**, *20*, 286–295. [[CrossRef](#)]
81. Dey, R.; Salem, F.M. Gate-variants of gated recurrent unit (GRU) neural networks. In Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA, 6–9 August 2017.
82. Wang, J.J.; Yan, J.; Li, C.; Gao, R.X.; Zhao, R. Deep heterogeneous GRU model for predictive analytics in smart manufacturing: Application to tool wear prediction. *Comput. Ind.* **2019**, *111*, 1–14. [[CrossRef](#)]
83. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [[CrossRef](#)]
84. Dhyani, M.; Kumar, R. An intelligent Chatbot using deep learning with Bidirectional RNN and attention model. *Mater. Today Proc.* **2021**, *34*, 817–824. [[CrossRef](#)]
85. Cui, Z.; Ke, R.; Pu, Z.; Wang, Y. Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values. *Transp. Res. Part C Emerg. Technol.* **2020**, *118*, 102674. [[CrossRef](#)]
86. Chen, D.; Yongchareon, S.; Lai, E.M.-K.; Yu, J.; Sheng, Q.Z.; Li, Y. Transformer with bidirectional GRU for nonintrusive, sensor-based activity recognition in a multiresident environment. *IEEE Internet Things J.* **2022**, *9*, 23716–23727. [[CrossRef](#)]
87. Mostafa, A.L.; Abdel-Galil, H.; Belal, M. Ensemble Model-based Weighted Categorical Cross-entropy Loss for Facial Expression Recognition. In Proceedings of the 2021 Tenth International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 5–6 December 2021.
88. Feng, L.; Shu, S.; Lin, Z.; Lv, F.; Li, L.; An, B. Can cross entropy loss be robust to label noise? In Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, Online, 7–15 January 2021.

89. Ruby, U.; Yendapalli, V. Binary cross entropy with deep learning technique for image classification. *Int. J. Adv. Trends Comput. Sci. Eng.* **2020**, *9*, 5393–5397.
90. Hernández-Vázquez, M.A.; Hernández-Rodríguez, Y.M.; Cortes-Rojas, F.D.; Bayareh-Mancilla, R.; Cigarroa-Mayorga, O.E. Hybrid Feature Mammogram Analysis: Detecting and Localizing Microcalcifications Combining Gabor, Prewitt, GLCM Features, and Top Hat Filtering Enhanced with CNN Architecture. *Diagnostics* **2024**, *14*, 1691. [[CrossRef](#)]
91. Suri, J.S.; Agarwal, S.; Carriero, A.; Paschè, A.; Danna, P.S.; Columbu, M.; Saba, L.; Viskovic, K.; Mehmedović, A.; Agarwal, S. COVLIA 1.0 vs. MedSeg: Artificial intelligence-based comparative study for automated COVID-19 computed tomography lung segmentation in Italian and Croatian Cohorts. *Diagnostics* **2021**, *11*, 2367. [[CrossRef](#)]
92. Daneshjou, R.; Smith, M.P.; Sun, M.D.; Rotemberg, V.; Zou, J. Lack of transparency and potential bias in artificial intelligence data sets and algorithms: A scoping review. *JAMA Dermatol.* **2021**, *157*, 1362–1369. [[CrossRef](#)] [[PubMed](#)]
93. Dubey, A.K.; Chabert, G.L.; Carriero, A.; Pasche, A.; Danna, P.S.; Agarwal, S.; Mohanty, L.; Nillmani; Sharma, N.; Yadav, S. Ensemble Deep Learning Derived from Transfer Learning for Classification of COVID-19 Patients on Hybrid Deep-Learning-Based Lung Segmentation: A Data Augmentation and Balancing Framework. *Diagnostics* **2023**, *13*, 1954. [[CrossRef](#)] [[PubMed](#)]
94. Saba, L.; Than, J.C.; Noor, N.M.; Rijal, O.M.; Kassim, R.M.; Yunus, A.; Ng, C.R.; Suri, J.S. Inter-observer variability analysis of automatic lung delineation in normal and disease patients. *J. Med. Syst.* **2016**, *40*, 1–18. [[CrossRef](#)]
95. Suri, J.S.; Agarwal, S.; Saba, L.; Chabert, G.L.; Carriero, A.; Paschè, A.; Danna, P.; Mehmedović, A.; Faa, G.; Jujaray, T. Multicenter study on COVID-19 lung computed tomography segmentation with varying glass ground opacities using unseen deep learning artificial intelligence paradigms: COVLIA 1.0 validation. *J. Med. Syst.* **2022**, *46*, 62. [[CrossRef](#)]
96. Unnikrishnan, P.; Kumar, D.K.; Arjunan, S.P.; Kumar, H.; Mitchell, P.; Kawasaki, R. Development of health parameter model for risk prediction of CVD using SVM. *Comput. Math. Methods Med.* **2016**, *2016*, 3016245. [[CrossRef](#)]
97. Zhou, R.; Guo, F.; Azarpazhooh, M.R.; Hashemi, S.; Cheng, X.; Spence, J.D.; Ding, M.; Fenster, A. Deep learning-based measurement of total plaque area in B-mode ultrasound images. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 2967–2977. [[CrossRef](#)]
98. Jain, P.K.; Sharma, N.; Giannopoulos, A.A.; Saba, L.; Nicolaidis, A.; Suri, J.S. Hybrid deep learning segmentation models for atherosclerotic plaque in internal carotid artery B-mode ultrasound. *Comput. Biol. Med.* **2021**, *136*, 104721. [[CrossRef](#)]
99. Jain, P.K.; Sharma, N.; Saba, L.; Paraskevas, K.I.; Kalra, M.K.; Johri, A.; Laird, J.R.; Nicolaidis, A.N.; Suri, J.S. Unseen artificial intelligence—Deep learning paradigm for segmentation of low atherosclerotic plaque in carotid ultrasound: A multicenter cardiovascular study. *Diagnostics* **2021**, *11*, 2257. [[CrossRef](#)] [[PubMed](#)]
100. Jain, P.K.; Sharma, N.; Saba, L.; Paraskevas, K.I.; Kalra, M.K.; Johri, A.; Nicolaidis, A.N.; Suri, J.S. Automated deep learning-based paradigm for high-risk plaque detection in B-mode common carotid ultrasound scans: An asymptomatic Japanese cohort study. *Int. Angiol.* **2021**, *41*, 9–23. [[CrossRef](#)] [[PubMed](#)]
101. Agarwal, M.; Agarwal, S.; Saba, L.; Chabert, G.L.; Gupta, S.; Carriero, A.; Pasche, A.; Danna, P.; Mehmedovic, A.; Faa, G. Eight pruning deep learning models for low storage and high-speed COVID-19 computed tomography lung segmentation and heatmap-based lesion localization: A multicenter study using COVLIA 2.0. *Comput. Biol. Med.* **2022**, *146*, 105571. [[CrossRef](#)]
102. Suri, J.S.; Bhagawati, M.; Paul, S.; Protogeron, A.; Sfrikakis, P.P.; Kitis, G.D.; Khanna, N.N.; Ruzsa, Z.; Sharma, A.M.; Saxena, S. Understanding the bias in machine learning systems for cardiovascular disease risk assessment: The first of its kind review. *Comput. Biol. Med.* **2022**, *142*, 105204. [[CrossRef](#)] [[PubMed](#)]
103. Narayanan, R.; Werahera, P.; Barqawi, A.; Crawford, E.; Shinohara, K.; Simoneau, A.; Suri, J. Adaptation of a 3D prostate cancer atlas for transrectal ultrasound guided target-specific biopsy. *Phys. Med. Biol.* **2008**, *53*, N397. [[CrossRef](#)]
104. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.
105. Bartlett, P.L.; Wegkamp, M.H. Classification with a Reject Option using a Hinge Loss. *J. Mach. Learn. Res.* **2008**, *9*, 1823–1840.
106. Bénédic, G.; Koops, V.; Odijk, D.; de Rijke, M. SigmoidF1: A smooth F1 score surrogate loss for multilabel classification. *arXiv* **2021**, arXiv:2108.10566.
107. Kyriacou, E.C.; Petroudi, S.; Pattichis, C.S.; Pattichis, M.S.; Griffin, M.; Kakkos, S.; Nicolaidis, A. Prediction of high-risk asymptomatic carotid plaques based on ultrasonic image features. *IEEE Trans. Inf. Technol. Biomed.* **2012**, *16*, 966–973. [[CrossRef](#)]
108. El-Barghouty, N.; Nicolaidis, A.; Bahal, V.; Geroulakos, G.; Androulakis, A. The identification of the high risk carotid plaque. *Eur. J. Vasc. Endovasc. Surg.* **1996**, *11*, 470–478. [[CrossRef](#)]
109. Stoitsis, J.; Golemati, S.; Nikita, K.; Nicolaidis, A. Characterization of carotid atherosclerosis based on motion and texture features and clustering using fuzzy c-means. In Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, CA, USA, 1–5 September 2004.
110. Loizou, C.P.; Petroudi, S.; Pantziaris, M.; Nicolaidis, A.N.; Pattichis, C.S. An integrated system for the segmentation of atherosclerotic carotid plaque ultrasound video. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control.* **2014**, *61*, 86–101. [[CrossRef](#)]
111. Stoitsis, J.; Golemati, S.; Kendros, S.; Nikita, K. Automated detection of the carotid artery wall in B-mode ultrasound images using active contours initialized by the Hough transform. In Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 20–25 August 2008.
112. Matsakou, A.I.; Golemati, S.; Stoitsis, J.S.; Nikita, K.S. Automated detection of the carotid artery wall in longitudinal B-mode images using active contours initialized by the Hough transform. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–2 September 2011.

113. Suri, J.S. Two-dimensional fast magnetic resonance brain segmentation. *IEEE Eng. Med. Biol. Mag.* **2001**, *20*, 84–95. [[CrossRef](#)] [[PubMed](#)]
114. Kiernan, M.J.; Al Mukaddim, R.; Mitchell, C.C.; Maybock, J.; Wilbrand, S.M.; Dempsey, R.J.; Varghese, T. Lumen segmentation using a Mask R-CNN in carotid arteries with stenotic atherosclerotic plaque. *Ultrasonics* **2024**, *137*, 107193. [[CrossRef](#)] [[PubMed](#)]
115. Zhou, R.; Guo, F.; Azarpazhooh, M.R.; Spence, J.D.; Gan, H.; Ding, M.; Fenster, A. Carotid vessel-wall-volume ultrasound measurement via a UNet++ ensemble algorithm trained on small data sets. *Ultrasound Med. Biol.* **2023**, *49*, 1031–1036. [[CrossRef](#)] [[PubMed](#)]
116. El-Baz, A.; Suri, J.S. *Big Data in Multimodal Medical Imaging*; CRC Press: Boca Raton, FL, USA, 2019.
117. Rumsfeld, J.S.; Joynt, K.E.; Maddox, T.M. Big data analytics to improve cardiovascular care: Promise and challenges. *Nat. Rev. Cardiol.* **2016**, *13*, 350–359. [[CrossRef](#)]
118. Krittanawong, C.; Johnson, K.W.; Hershman, S.G.; Tang, W.W. Big data, artificial intelligence, and cardiovascular precision medicine. *Expert Rev. Precis. Med. Drug Dev.* **2018**, *3*, 305–317. [[CrossRef](#)]
119. Hulsen, T.; Friedecký, D.; Renz, H.; Melis, E.; Vermeersch, P.; Fernandez-Calle, P. From big data to better patient outcomes. *Clin. Chem. Lab. Med. CCLM* **2023**, *61*, 580–586. [[CrossRef](#)] [[PubMed](#)]
120. Dabla, P.K. Unlocking new potential of clinical diagnosis with artificial intelligence: Finding new patterns of clinical and lab data. *World J. Diabetes* **2024**, *15*, 308. [[CrossRef](#)]
121. Wang, S.; Xie, T.; Liu, H.; Zhang, X.; Cheng, J. PSE-Net: Channel pruning for Convolutional Neural Networks with parallel-subnets estimator. *Neural Netw.* **2024**, *174*, 106263. [[CrossRef](#)]
122. Louati, H.; Louati, A.; Bechikh, S.; Kariri, E. Embedding channel pruning within the CNN architecture design using a bi-level evolutionary approach. *J. Supercomput.* **2023**, *79*, 16118–16151. [[CrossRef](#)]
123. Hong, W.; Li, G.; Liu, S.; Yang, P.; Tang, K. Multi-objective evolutionary optimization for hardware-aware neural network pruning. *Fundam. Res.* **2022**, *4*, 941–950. [[CrossRef](#)] [[PubMed](#)]
124. Hollmann, N.; Müller, S.; Eggensperger, K.; Hutter, F. TabPFN: A transformer that solves small tabular classification problems in a second. *arXiv* **2022**, arXiv:2207.01848.
125. Huang, X.; Khetan, A.; Cvitkovic, M.; Karnin, Z. Tabtransformer: Tabular data modeling using contextual embeddings. *arXiv* **2020**, arXiv:2012.06678.
126. Somepalli, G.; Goldblum, M.; Schwarzschild, A.; Bruss, C.B.; Goldstein, T. Saint: Improved neural networks for tabular data via row attention and contrastive pre-training. *arXiv* **2021**, arXiv:2106.01342.
127. Makris, G.C.; Lavida, A.; Griffin, M.; Geroulakos, G.; Nicolaides, A.N. Three-dimensional ultrasound imaging for the evaluation of carotid atherosclerosis. *Atherosclerosis* **2011**, *219*, 377–383. [[CrossRef](#)]
128. Kyriacou, E.C.; Pattichis, C.; Pattichis, M.; Loizou, C.; Christodoulou, C.; Kakkos, S.K.; Nicolaides, A. A review of noninvasive ultrasound image processing methods in the analysis of carotid plaque morphology for the assessment of stroke risk. *IEEE Trans. Inf. Technol. Biomed.* **2010**, *14*, 1027–1038. [[CrossRef](#)]
129. Chen, Z.; Jiang, M.; Chiu, B. Unsupervised shape-and-texture-based generative adversarial tuning of pre-trained networks for carotid segmentation from 3D ultrasound images. *Med. Phys.* **2024**. *early view*. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.