A Hybrid Deep Learning Technique for Feature Selection and Classification of Chronic Kidney Disease

Ramya Asa Latha Busi[1]   James Stephen Meka[2]*   P V G D Prasad Reddy[1]

[1]*Department of Computer Science & Software Engineering, Andhra University, Visakhapatnam, India*
[2]*Dr. B. R. Ambedkar Chair, Andhra University, Visakhapatnam, India*
* Corresponding author's Email: stephenprofmjames@gmail.com

Abstract: A significant global public health concern is the widespread presence of chronic kidney disease (CKD). High mortality rates are associated with the disease, particularly in developing nations. Since there are no visible early-stage signs, CKD frequently goes undiagnosed. In the meantime, preventing the disease from progressing requires early detection and prompt clinical care. To help clinicians discover CKD early, deep learning (DL) techniques can give an effective and affordable diagnosis. This research proposes a unique hybrid DL approach to classify CKD. In a pre-processing step, eliminate the missing values and reduce noise from data, data transformation, and outlier detection. After that, using the improved capsule network (Improved CapsNet) method to extract the features. Then, select essential features using the improved spotted hyena optimizer (ISHO) algorithm to better classification with less time. Finally, employ hybrid deep learning techniques of BConvLSTM and DNetCNN to classify the CKD. A recently introduced CKD prediction algorithm and well-known classifiers were used as benchmarks for the proposed approaches. The proposed model, which was trained with the smaller feature set, outperformed other classifiers with a classification accuracy rate of 99.89%. The experimental findings also demonstrate the positive effect of feature selection on the performance of the different techniques. The proposed technique has developed a reliable predictive system for recognizing CKD and may be extended to more unbalanced medical datasets to identify diseases reliably.

Keywords: Chronic kidney disease, BConvLSTM, DNetCNN, Improved spotted hyena optimizer (ISHO) algorithm, Improved capsnet, Deep learning.

Notations

S_j	Capsule Input
v_j	Capsule Output
W_{ij}	Weight Matrix
$U_{j/i}$	Predictive Vector
a_{ij}	Log-Likelihood
X	Current Iteration
$\vec{A}\vec{E}$	Co-Efficient Vectors
$\overrightarrow{P_q(x)}$	Prey's Location Vector
$\overrightarrow{P_q(x)}$	Spotted Hyena's Location Vector

1. Introduction

The high mortality rate of CKD has prompted a lot of attention. According to the world health organization (WHO), CKD has grown to be a problem that threatens developing nations. CKD is curable in its early stages but develops kidney failure in its last stages [1, 2]. Around the world, chronic kidney disease killed 800 million lives in 2022. The reason kidney illness is defined as a "chronic" condition is that it develops slowly over time and has an impact on the urinary system. Other health issues that have a variety of symptoms, including diabetes, low and high blood pressure, bone issues, and nerve damage, which result in cardiovascular disease, are brought on by the buildup of waste products in the blood [3-5]. Diabetes, cardiovascular disease (CVD), and blood pressure are challenging impacts for CKD-affected people. Patients may develop the disease in advanced phases in developing countries through kidney transplantation or necessitating dialysis [6, 7].

CKD can be protected from kidney failure with early detection and treatment. The easiest way to control CKD is to recognize it early, but waiting until it has progressed too far will result in kidney damage and the need for continuous dialysis or a kidney transplant to maintain everyday life [8–10]. Two medical procedures diagnose chronic kidney disease: blood or urine tests. CAD needs to support the diagnostic decisions made by doctors and radiologists because of the growing amount of patients with CKD, the shortage of specialists, and the high amount of treatment and diagnosis, particularly in developing nations [11, 12]. When implementing ML tasks, the techniques use the discriminative properties of the attributes to categorize the samples. The effectiveness of ML techniques depends not only on the selected technique but also on the characteristics of the input information [13-15]. Furthermore, not all input features may be equally significant in most ML applications, particularly in medical diagnosis. This lowers the computing cost of the model construction. To improve the proposed method's performance, we hybrid the novel deep learning techniques to classify chronic kidney diseases. The feature selection chooses necessary attributes to get less computation time. Hybrid deep learning techniques classify CKD. The main essential contribution of this research is,

- Missing variables, data transformation, encoding, and outlier recognition are done in a pre-processing step. To extract the features, use Improved CapsNet to extract the features from the dataset.
- Select the necessary attributes from extracted attributes utilizing the Improved Spotted Hyena Optimization (ISHO) algorithm to classify the CKD less quickly.
- To classify the chronic kidney disease into whether it's CKD or not, employed hybrid techniques of BConvLSTM and DNetCNN.
- The evaluation performance is done with the CKD dataset and their attributes with accuracy, recall, specificity, precision, and f-measure metrics.

The other parts of the paper are arranged as follows. Section 2 discusses the related study on chronic kidney disease classification. Section 3 discusses in extensive detail the proposed technique and its parts. Section 4 describes the experimental approach. In section 5, the work is discussed, along with ideas for further research.

2. Literature survey

In this section, we mentioned some previous research papers based on CKD. Senan et al. [16] recommended using efficient categorization algorithms and recursive feature elimination methods to recognize CKD. In the pre-processing stage, they estimate missing variables and eliminate noise, such as normalization and outliers. Then, the recursive feature elimination (RFE) algorithm was used to identify the essential attributes by finding a high level of correlation between particular attributes and the targets. After that, the classification of the CKD utilizes four ML techniques (SVM, KNN, DT, RF).

A diagnostic recognition model for CKD was suggested by Hosseinzadeh et al. [17]. In this paper, they collected necessary data with smart multimedia medical devices and biomedical sensors. In a pre-processing step, they cleaned the collected data from any unnecessary noises and inconsistencies for the training and testing step in categorization performance. Then, select essential features to improve the effectiveness of learning algorithms. It finally used four different ML techniques, such as SVM, MLP, DT, and Naive Bayes, to classify chronic kidney diseases.

Early forecasting of CKD utilizing the deep belief network method was introduced by Elkholy et al. [18]. This paper introduced an intelligent categorization and forecasting model. First, they collect data from the UCI Dataset. After that, they remove the missing variables, reducing efficiency before analyzing the data. Finally, they used a modified deep belief network (DBN) as a categorization approach to forecasting kidney-related illness and the categorical cross-entropy as a loss function, and the Softmax as an activation function.

For the diagnosis and categorization of CKD, Elhoseny et al. [19] suggested the density-based feature selection (DFS) with ant colony based optimization (D-ACO) approach. The pre-processing step is the initial process since the database may contain noise and redundant information. Examining the information allows for several processes to be carried out, such as data cleansing, filling in for missing numbers, and deleting extraneous data, as both of these affect performance. There are 24 features in total in this work, some of which are chosen using DFS by repeatedly producing a set of attributes until DFS obtains the best subset; a wrapper approach was used to select the best feature subset. ACO-based categorization approach is used to categorize the data as to whether there is CKD or not having CKD to register the obtained feature vector.
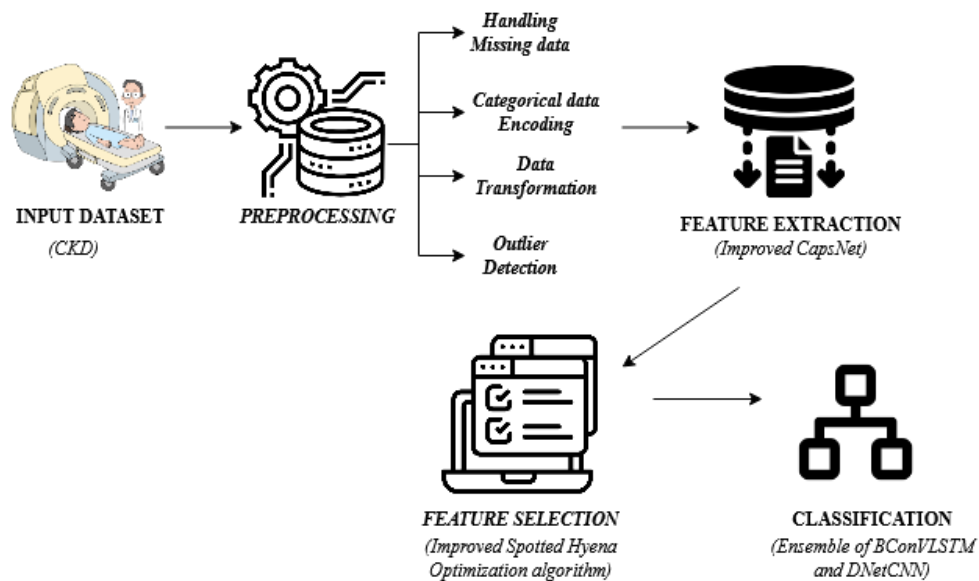
Figure. 1 The structure of the proposed work

To predict CKD based on cloud IoT, Abdelaziz et al. [20] suggested a hybrid intelligent model employing linear regression (LR) and neural network (NN), two intelligent techniques. Using LR, it is possible to identify significant variables that affect CKD. NN was employed to forecast CKD.

From the literature survey, they have some research problems to solve. The problems occur in literature studies,

- They did not perform topic modeling and accuracy is not significant.
- The machine learning algorithms might need more interpretability, making it easier for healthcare professionals to understand and trust the decision-making process.
- They didn't select any relevant features for CKD detection.
- The computation complexity of the research is high.

To overcome these problems, propose a novel deep-learning technique to achieve effective results in this research. A process on dataset data to estimate the missing values, remove noise, data transformation, and outlier detection is done in a pre-processing step. Then, select the essential features from extracted features to perform classification accurately with less computation time.

3. Proposed methodology

CKD has become a global burden, increasing mortality rates due to the lack of an adequate diagnosis system. Providing improved medical care for CKD patients while preserving people's lives is becoming more difficult these days. To address the CKD categorization problem, an efficient hybrid BConvLSTM and DNetCNN technique is developed in this research. The proposed technique involves the four essential steps previously discussed once the dataset was collected. It pre-processes handling missing data, transformation, categorial data encoding, and outlier recognition. Feature extraction is utilized to retrieve the features from the dataset. Feature selection is utilized to choose the necessary attributes. Categorization of CKD for improving recognition accuracy. The CKD information is pre-processed, then the attributes are extracted based on the improved CapsNet approach. And then, the attributes are chosen based on an improved spotted hyena optimization algorithm. Finally, the ensemble approach is proposed to classify the attributes into CKD or Non-CKD.

Fig. 1 demonstrates the structure of the proposed work.

3.1 Pre-processing

The pre-processing processes included the missing variables estimation and the elimination of noise. Some measurements during a patient evaluation could be absent or insufficient.

Handling missing values

In the data collection, 158 cases are complete, while the remaining ones are missing. The simplest method for handling missing variables is to ignore records, although this is impractical for small data

641

sets. The data set is checked to see whether any attribute variables are missing. The statistical approach of mean computation was utilized to approximate the missing variables for the numerical attributes. The missing values of nominal characteristics were replaced using the mode approach.

Categorical data encoding

Category variables must be encoded into number variables because most DL techniques only accept numeric variables as input. The features of categories like "no" and "yes" are represented by the binary values "0" and "1".

Data transformation

The process of altering amounts on the same scale is known as data transformation, which is done to prevent one variable from overpowering the others. Otherwise, regardless of the unit of weight, learning algorithms interpret more significant variables as smaller and higher ones as lower. To enable additional processing, data transformations change the variables of a data set. The research involved uses a data normalization strategy to increase the accuracy of DL models. The transformed data has a mean of 0 and a standard deviation 1.

Outlier detection

Observational points that stand out from the rest of the data are known as outliers. Measurement variability may be the root cause of an outlier, or it may indicate an experimentation error. The deep learning algorithm's learning process can be distorted and misled by an outlier. It results in longer training times, lower model accuracy, and worse outcomes. This research employs the interquartile range (IQR)-based technique to eliminate outliers.

3.2 Feature extraction

After performing pre-processing, we used the CapsNet technique to extract the features from the dataset. The Improved CapsNet model, a feature extractor, is applied for improved classification performance. When attributes or variables are actively processed using a variable or attribute selection mechanism, the FV outcomes may have redundant or unneeded attributes. The entire performance of choosing a portion of the highly discriminant characteristics is called feature selection. The proposed research employs entropy to assess

uncertain information and show signal unpredictability by exhibiting the model disorder.

An established CapsNet method describes hierarchical relationships and maintains object position and feature information in the data. The pooling layer receives helpful information from the information in the CNN method. The likelihood of the network learning minute details decreases when the information is transmitted to the next pooling layer. The neural result of the CNN approach also produces a scalar value. The CapsNets generate vector results of a similar size but with different routing; each capsule comprises multiple neurons. The variable of data is denoted by the vector route [22]. CapsNet uses the Eq. (1) specified vector activation function called squashing as an alternative.

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2} \frac{s_j}{\|s_j\|} \qquad (1)$$

j and S_j stand for the total capsule input, and v_j stands for capsule output. When there is an item in the data, v_j it shrinks the long vector to one, and when there isn't, it chokes the shorter vector to zero.

In addition to the initial layer of CapsNet, the weighted amount of the forecasted vector $\left(U_{j/i}\right)$ in the capsule, as located in the lower layer, estimates the total input values of the capsule S_j. A capsule from the lowest layer's outcome (O_i) and weight matrix (W_{ij}) are used to estimate the predictive vector $\left(U_{j/i}\right)$.

$$S_j = \sum_i b_{ij} u_{j/i} \qquad (2)$$

$$u_{j/i} = W_{ij} O_i \qquad (3)$$

Where b_{ij} stands for the coefficient determined by the dynamic routing technique and is determined by,

$$b_{ij} = \frac{\exp(a_{ij})}{\sum_k \exp(a_{ik})} \qquad (4)$$

Where log-likelihood is denoted by a_{ij}. One is the correlation coefficient between capsules i and those in the top layer, while Softmax defines log prior likelihood. For determining the objects of a specific class that are present in CapsNet, a margin loss is offered and is evaluated as follows:

$$L_k = T_k max[O, m^+ - \|v_k\|]^2 \\ + \lambda[1 - T_k max[O, \|v_k\| - m^-]^2] \qquad (5)$$

When class k is present, the value of T_k is one. Additionally, $m^+ = 0.9$ and $m^- = 0.1$ represent the weight of the loss and the hypervariable, respectively.

3.3 Feature selection

After extracting the features, it is necessary to find the essential components that have a strong and positive connection with attributes of importance for disease recognition. A robust recognition technique cannot be built since the vector features must be retrieved to exclude attributes that could be more helpful and relevant for forecasting. The improved spotted hyena optimizer (ISHO) algorithm is well known for its simplicity of use, adaptability to various settings, and efficiency in picking features from training datasets pertinent to forecasting target variables and removing weak attributes. The ISHO approach chooses the essential attributes by identifying high correlations between particular attributes and the target.

3.3.1. Improved spotted hyena optimizer (ISHO)

Large carnivorous canines, known as spotted hyenas, can be found in various open, arid habitats. Spotted hyenas feast on large and medium-sized herbivores, including wildebeests, impalas, and zebras. The spotted hyena is a highly sociable and clever animal. They use a variety of senses to recognize relatives and other people [23]. The connections between people of the same race were also ranked. In a population, trust is prioritized for those with high status.

3.3.1.1. Encircling prey

Spotted hyenas can locate their prey and encircle them. Due to the unknown search space, the spotted hyena closest to the prey is now the best contender. Following the determination of the optimum search solution, the locations of each other search agents are updated.

The below equations represent the mathematical description of this behavior:

$$C_h = \left| \vec{A} . \overrightarrow{P_q(z)} - \overrightarrow{P(z)} \right| \tag{6}$$

$$\overrightarrow{P(z+1)} = \overrightarrow{P_q(z)} - \vec{E} \cdot \overrightarrow{D_h} \tag{7}$$

Following is a computation of the vectors B and E:

$$\vec{A} = 2.r\overrightarrow{d_1} \tag{8}$$

$$\vec{E} = 2\vec{h} \cdot r\overrightarrow{d_2} - \vec{h} \tag{9}$$

$$\vec{h} = 5 - \left[iteration \frac{5}{MAX_{iteration}} \right] \tag{10}$$

3.3.1.2. Hunting

Spotted hyenas usually hunt in packs, rely on a group of reliable allies, and are good at spotting their prey. To update their location, other search agents should move in the direction of the best search agent. In this mechanism, the following equations are provided:

$$\overrightarrow{C_h} = \left| \vec{A} \cdot \overrightarrow{P_h} - \overrightarrow{P_k} \right| \tag{11}$$

$$\overrightarrow{P_k} = \overrightarrow{P_h} - \vec{E} \cdot \overrightarrow{C_h} \tag{12}$$

$$\overrightarrow{S_h} = \overrightarrow{P_k} + \overrightarrow{P_{k+1}} + \cdots + \overrightarrow{P_{k+N}} \tag{13}$$

The letters define the $\overrightarrow{P_{k+1}}$ $\overrightarrow{P_h}$ other spotted hyenas, which designate the location of the first best-spotted hyena. Here, N stands for the number of spotted hyenas, and its value is calculated as follows:

$$N = count_{nos}\left(\overrightarrow{P_h}, \overrightarrow{P_{h+1}}, \overrightarrow{P_{h+2}}, \cdots, \overrightarrow{P_h} + \vec{M} \right) \tag{14}$$

While nos specify the number of options and calculate all candidate options that are very similar when added to the best optimal answer found during the search, a random vector named $\vec{M}$ has a value between [0.5, 1].

3.3.1.3. Attacking prey (Exploitation)

The vector's value is reduced to create a mathematical model for attacking the target. To adjust the vector's value, which throughout iterations can go from 5 to 0, the variation in the vector is also reduced. When $E_j > 1$, a pack of spotted hyenas charges its victim. The following is the formula for attacking the prey mathematically:

$$\overrightarrow{P_{(z+1)}} = \frac{\overrightarrow{C_h}}{N} \tag{15}$$

Where the best search agent is located, other search agents are updated to reflect that location, and the best solution is recorded. With the help of the SHO approach, search agents can alter their locations and move closer to the target.

3.3.1.4. Search for prey (Exploration)

The position of the spotted hyenas in vector $\overrightarrow{C_h}$ spotted hyenas mostly hunt for prey. They avoid one another so they can hunt and attack their prey. When $E_j > 1$, spotted hyenas should depart from their prey. The system enables SHO to conduct international searches. A further element of SHO is $\vec{A}$ it enables exploration. The $\vec{A}$ vector in Eq. (15) contains the random value determining the prey's weight.

3.4 Classification

A hybrid deep learning technique is used to classify CKD. Here, we employed Bi-Directional Convolutional Long-Short Term Memory and DarkNet convolutional neural network. The data's global, long-term spatiotemporal characteristics are not altered when a ConvLSTM layer is added to the intermediate representations' spatial scale. Throughout the LSTM's recurrent operation, this encoding requires performed. A learned temporal representation is produced by a conventional LSTM network, which vectorizes and encodes its input through fully connected layers. These fully connected layers cause a loss of spatial information. Therefore, a convolutional rather than a fully linked operation may be preferred to maintain such spatial information. The ConvLSTM achieves that. It substitutes convolutional layers for the fully connected layers in the LSTM.

3.4.1. Bi-directional convolutional long-short-term memory (BConvLSTM)

Bi-Directional Convolutional Long-Short Term Memory (Bi-C-LSTM) is an LSTM neural network architecture variant incorporating convolutional layers and bidirectional processing. It combines the strengths of CNNs and LSTMs to capture spatial and temporal dependencies in sequential data. An extended recurrent neural network (RNN) is an LSTM, which can better consider time dependency [24]. Because LSTM cannot incorporate spatial correlation, ConvLSTM can address this issue. Both input-to-state and state-to-state transitions are subject to convolution procedures. ConvLSTM has a forget gate, an output gate, a memory cell, and an input gate, the same as LSTM. ConvLSTM can be expressed in the following way:

$$i_t = \sigma\left(A_{x_i} X_t + A_{h_i} H_{t-1} + A_{c_i} C_{t-1} + b_i\right) \quad (16)$$

$$f_t = \sigma\left(A_{x_f} X_t + A_{h_f} H_{t-1} + A_{c_f} C_{t-1} + b_f\right) \quad (17)$$

$$C_t = f_t o C_{t-1} + i_t tanh\left(A_{x_c} * X_t + A_{h_c} * H_{t-1} + b_c\right) \quad (18)$$

$$o_t = \sigma\left(A_{x_o} X_t + A_{h_o} H_{t-1} + A_{c_o} o C_t + b_c\right) \quad (19)$$

$$H_t = o_t o tanh(C_t) \quad (20)$$

The convolution operation is represented by $*$, and Hadamard product by o. In contrast to ConvLSTM, which only employs forward data dependencies, BConvLSTM utilizes both backward and forward ConvLSTMs to perform the input data. The backward and forward information dependence can enhance the performance of predictions.

$$Y_t = tanh\left[A_y^{\overrightarrow{H}} \cdot \overrightarrow{H} + A_y^{\overleftarrow{H}} \cdot \overleftarrow{H} + b\right] \quad (21)$$

Tanh is a nonlinear function that is utilized to combine the result.

3.4.2. Darknet convolutional neural network (DNetCNN)

After selecting the essential features, we employ a new deep learning technique DNetCNN to categorize the CKD into whether it's CKD or non-CKD. CNN is the best classifier among the many DL techniques. To increase the classification accuracy of the classifying CKD, the DNetCNN framework is utilized in this research. The activation function activates five pooling layers, 19 convolutional layers, and the darknet structure as an entire entity. The function of sigmoid activation is utilized for binary categorization [25]. The Softmax activation function is utilized in multi-classification. The 2D convolutional process is carried out with Eq. (22) for the input data X and kernel K.

$$C(X,K)_{(i,j)} = \sum_r \sum_c K(r,c) \times X \times (i-r, j-c) \quad (22)$$

Step parameterized input matrix for K. Given that the testing dataset for CKD uses a binary classification with Eq. (23), the sigmoid function is utilized in the proposed DNetCNN as an activation function.

$$Sigmoid(h) = \frac{1}{1+e^{-X_i}} \quad (23)$$

There were 16 convolution layers in the proposed DNetCNN. One convolutional layer comprising the activation and convolution procedures was present in

every darknet layer. The follow-up of the three successive forms was the same for every of the four convolution layers. Standardizing the input data due to the convolution layer's typical operation reduced training time. The Maxpool operation in the pooling layer was carried out using the 2*2 formula. The region that the filters employed will be maximized. The filter sizes ranged from 8, 16, 32, and 64 from the darknet to the convolution layer. Fig. 3 depicts the proposed DNetCNN's layered architecture. The first layer was the darknet layer with a 3*8 filter. Then followed DN, pool, then CNN in that order. The filter value of the last convolution layer was 256. Using the DNetCNN technique classified CKD accurately. It achieved higher classification accuracy with less computation time.

Pseudo-code

```
BEGIN
1.  Importing a CKD data set
2.  Fill in the blanks and eliminate outliers
3.  Convert text into numerical values
4.  Adjust the data scale
5.  Make use of the ISHO 'Feature Selection
6.  FOR FEATURE SELECTION METHODS:
7.  Choose the most essential features.
8.  Make a list of the essential features.
9.  Count the number of times each feature
    appears.
END FOR
10. Provide the proposed model with a list of
    features.
11. Set the model hyperparameters
12. Assign different validation scores to the
    scoring
13. Train using the proposed model
14. Sort data into CKD and non-CKD categories
15. Validation score
END
```

4. Results and discussion

In this part, the results of the proposed approach are reported. In Table 1, the hyperparameter settings for the proposed models are displayed. Figure 8 displays the confusion matrices. It proves the provided model successfully classified all actual positive and actual adverse events. The CKD class reports recall, accuracy, f-measure, and precision.

4.1 Experimental setup

Different environments have been utilized in the

Table 1. Settings for hyper-parameters

Hyper-Parameter	Setting
Batch size	15
Epochs	850
Activation Function	Relu
Optimizer	Adam
Dropout rate	0.5 to 0.1
Loss	Binary_Crossentropy
Activation output layer	Sigmoid

Table 2. The proposed system's environment setup

Resource	Details
RAM	8 GB
CPU	Core i5 Gen6
Software	Python
GPU	4 GB

Table 3. Splitting dataset

Dataset	Numbers
Testing and validation	100 patients
Training	300 patients

system's development. The environment configuration for the evolving system is displayed in Table 2.

4.2 Dataset description

Data on CKD were gathered via the university of california irvine (UCI) Repository. The data set contains 400 patient records, some needing more specific values. It consists of 24 clinical attributes that show up in the prognostic of CKD, with a class attribute acting as an outcome of the patient being expected to have CKD. The expected attribute diagnostic has two values: "ckd" and "non-ckd." One hundred fifty values from the "non-ckd" class (37.5%) and 250 values from the "ckd" class (62.5%) are both included in the data set. (Source: https://www.kaggle.com/datasets/mansoordaku/ckdisease)

Splitting dataset: The dataset was split into 25 for testing and validation and 75% for training. The split data are displayed in Table 3.

4.3 Evaluation metrics

The accuracy of the proposed approach was computed by setting the non-CKD class value to zero and the CKD class value to one. True positives asserts that CKD has been correctly classified. The results of the false negatives test demonstrate that CKD was miscategorized. A false-positive result (FP) shows that the non-CKD samples were not correctly detected. True negative (TN) samples were correctly labeled as not having CKD.
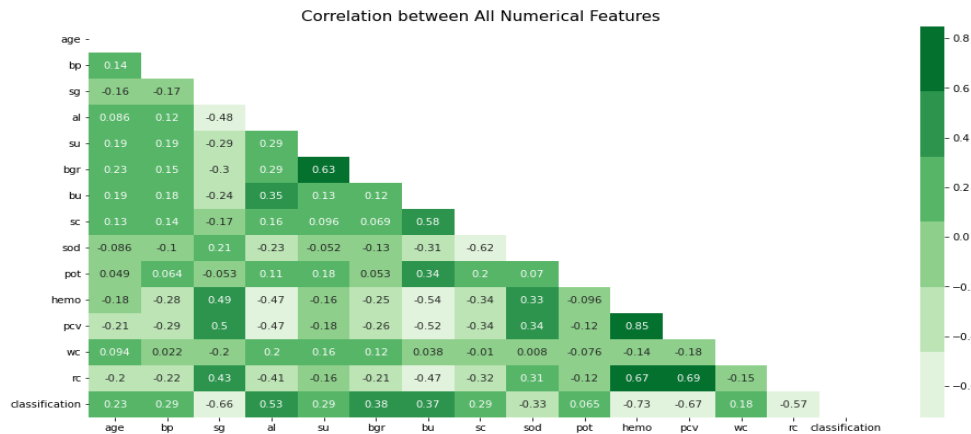
Figure. 2 Correlation between several features

4.3.1. Accuracy

It is the ratio of accurate estimates to all forecasting. The capacity to predict outcomes accurately can be used to define accuracy.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (24)$$

4.3.2. Recall

The following equation demonstrates how recall determines the ratio of correctly anticipated positive observations.

$$Recall = \frac{TP}{TP+FN} \qquad (25)$$

4.3.3. Precision

This metric denotes the ratio of correctly forecasted positive observations to all optimistic predictions, as shown in the equation below.

$$Precision = \frac{TP}{TP+FP} \qquad (26)$$

4.3.4. F-measure

In the F-measure, recall and precision are weighted and averaged. The method involves both false negatives and false positives. The term "F-Measure" is defined as

$$F - Measure = \frac{2 \times (Precision \times Recall)}{(Precision+Recall)} \qquad (27)$$

The range of F-measure values is 0 to 1.

4.4 Performance metrics

The performance of the proposed method is to classify chronic kidney disease accurately with a higher classification accuracy level. Correlation is a statistical measure that quantifies the connection between two or more attributes. It indicates how changes in one attribute are combined with changes in another. The correlation coefficient, often defined as "r," ranges from -1 to 1, where 0 indicates no correlation, 1 indicates a strong positive correlation, and -1 indicates a strong negative correlation.

A dataset can be used to calculate the correlation between different traits or variables. The degree and direction of the association between features can be understood using correlation analysis. A positive correlation indicates that the two variables rise as one increases. When two variables are negatively correlated, one variable tends to increase while the other tends to decline. The absence of correlation indicates no consistent connection between the variables. Fig. 2 illustrates the adjustment between several dataset features.

4.5 Feature selection results

The proposed feature selection approaches are compared with other approaches in this section. Accuracy, recall, and specificity are used as evaluation criteria. Previous feature selection techniques are utilized, such as CFS+SVM, MIFS+SVM, RFS+SVM, Relief + CSF+ SVM, and RFP+SVM. The comparison of feature selection methods is shown in Table 4.

Table 5 compares existing and proposed classification techniques for CKD dataset analysis. Compared to other techniques, the proposed method achieved higher accuracy values, such as 99.89% of accuracy, 100% of recall, 99.85% of precision, and 99.82% of f-measure. Still, the existing method SVM obtained lower values of 92% of accuracy, 87% f recall, 96% of precision, and 92% of f-measure.

Table 4. Comparison of feature selection approaches

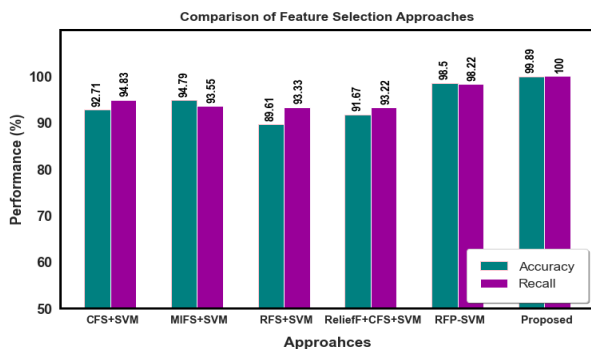| Feature Selection Technique	Total no of features	Selected features	% of features eliminated	Acc (%)	Rc (%)
Correlation feature selection +SVM	22	11	50	92.71	94.83
Mutual information-based Feature selection +SVM	22	12	45.45	94.79	93.55
Relief Feature selection +SVM	22	8	63.63	89.61	93.33
ReliefF +CFS+SVM	22	7	68.18	91.67	93.22
R_FP-SVM	22	7	68.18	98.50	98.22
Proposed	22	10	70	99.89	100

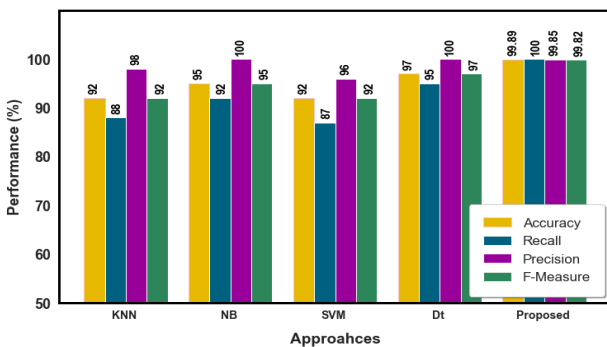Figure. 3 Graphical representation of feature selection approaches

Figure. 4 Comparative analysis of classification techniques of proposed and existing techniques

Table 5. Comparison of the proposed model to the dataset for CKD's existing classification methods

Method	Accuracy (%)	Recall (%)	Precision (%)	F-Measure (%)
K-Nearest Neighbor (KNN)	92	88	98	92
Naïve Bayes (NB)	95	92	100	95
Support Vector Machine (SVM)	92	87	96	92
Decision Tree (DT)	97	95	100	97
Proposed	99.89	100	99.85	99.82

Table 6. Comparison of the proposed technique with other techniques from the literature used with the UCI dataset

Authors	Model	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)
Senan et al. [16]	SVM	97.3	94.74	92	96.67
Hosseinzadeh et al. [17]	DT	97	95	99	-
Elkholy et al. [18]	DBN	98.52	-	-	87
Elhoseny et al. [19]	D-ACO	95	93.33	96	96
Abdelaziz et al. [20]	LR+NN	97.8	96.2	100	98.1
Proposed	BConvLSTM +DNetCNN	99.89	99.85	100	99.82

The comparative analysis of the existing and proposed methods is shown in Fig. 4 as a graph representation.

We compare various approaches with our proposed method. Table 6 shows how our proposed technique achieved the highest effectiveness and increased the detection rate. The proposed technique operates accurately with the dataset.

Table 6 lists the results for SVM, DT, DBN, D-ACO, AlexNet, LR+NN, and the proposed BConvLSTM+DNetCNN on the Dataset in terms of accuracy. The proposed technique achieved higher categorization accuracy than others.
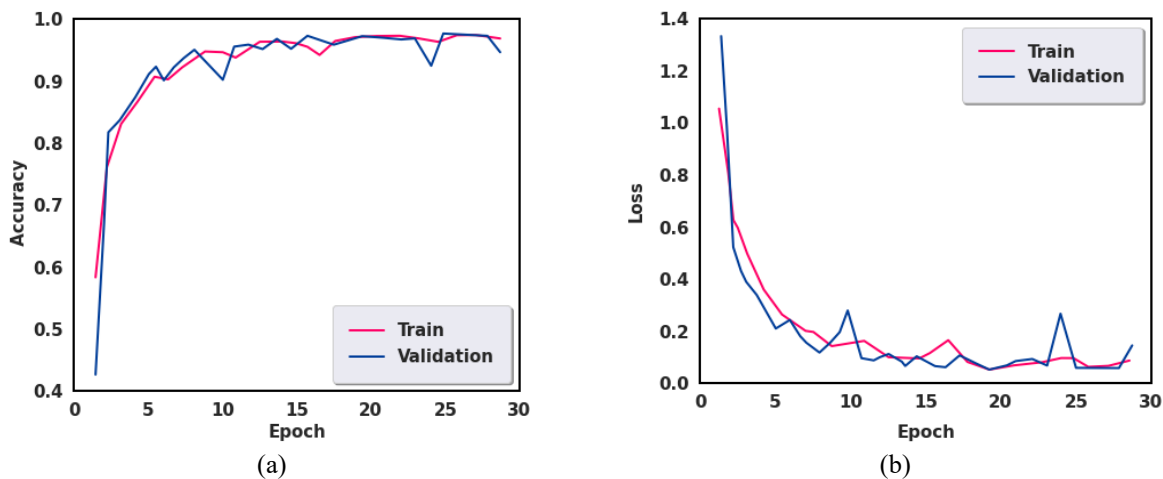
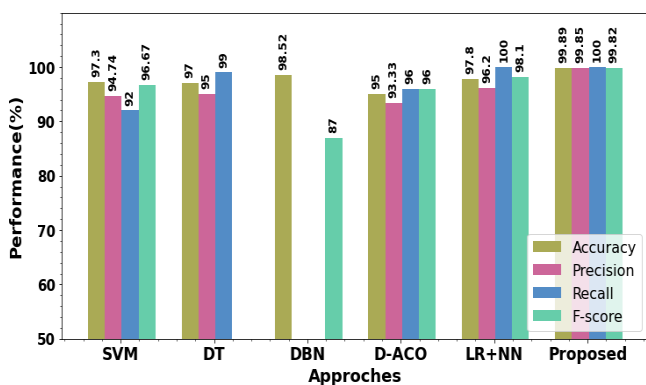Figure. 6: (a) Accuracy for testing and training and (b) Testing and training loss for the UCI CKD dataset

Figure. 5 Comparison of proposed with previous techniques

Table 7. Utilizing the proposed and previous methods

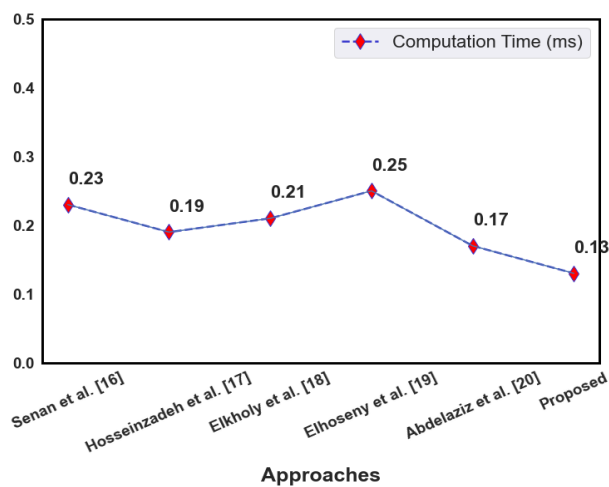| Methods	Computation Time
Senan et al. [16]	0.23
Hosseinzadeh et al. [17]	0.19
Elkholy et al. [18]	0.21
Elhoseny et al. [19]	0.25
Abdelaziz et al. [20]	0.17
Proposed	0.14

Figure. 7 The computation time of proposed and previous techniques

The achieved accuracy for the proposed technique is 99.89%, compared to SVM at 96.67%, DT at 97%, DBN at 98.5%, D-ACO at 95%, and LR+NN at 97.8%. Fig. 5 shows the comparison of the proposed approach with the existing approach.

4.6 Evaluation of training and testing

As the number of iteration steps increased, a graph of loss value and classification accuracy is shown in Fig. 6. The graph demonstrates that the method discussed in this study positively impacts convergence. The dataset was divided into testing and training phases. The processed training set is used in the training phase to train the proposed methods for 200 iterations. The learning rate is currently at 0.1.

4.7 Computation time

The error and computational time of the selected classifier serve as performance indicators. For classification accuracy, precision and recall will be considered. Each classifier takes into account the computing time. The proposed BConvLSTM+DNetCNN method achieves good recall, f-measure, and accuracy with low computation time. The proposed BConvLSTM+DNetCNN system's computing time is shown in Fig. 7.

Our proposed method achieved higher prediction accuracy and less computation time than previous techniques.

5. Conclusion and future scope

Chronic kidney disease (CKD) is a long-term condition in which the kidneys cannot function properly. It is a progressive condition that worsens over time, often leading to permanent kidney damage and declining kidney function. In this work, we

proposed a unique hybrid DL approach to classify CKD. In a pre-processing step, eliminate the missing values and reduce noise from data, data transformation, and outlier detection. After that, it uses the improved capsule network (Improved CapsNet) method to extract the features. Then, select essential features using the improved spotted hyena optimizer (ISHO) algorithm to better classification with less time. Finally, employ hybrid deep learning techniques of BConvLSTM and DNetCNN to classify the CKD. The performance of the proposed model achieved 99.89% accuracy with less computation time. The future scope will develop a hybrid technique with an optimization algorithm to increase the accuracy of disease identification before the condition reveals itself in humans.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions

Conceptualization, Ramya Asa Latha Busi and James Stephen Meka; methodology, P V G D Prasad Reddy; software, Ramya Asa Latha Busi; validation, Ramya Asa Latha Busi, James Stephen Meka and P V G D Prasad Reddy; formal analysis, James Stephen Meka; investigation, P V G D Prasad; resources, James Stephen Meka; writing—original draft preparation, Ramya Asa Latha Busi; writing—review and editing, P V G D Prasad.

Acknowledgment

We declare that this manuscript is original, has not been published before, and is not currently being considered for publication elsewhere.

References

[1] F. Ma, T. Sun, L. Liu, and H. Jing, "Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network", *Future Generation Computer Systems*, Vol. 111, pp. 17-26, 2020.

[2] V. Singh, V. K. Asari, and R. Rajasekaran, "A deep neural network for early detection and prediction of chronic kidney disease", *Diagnostics*, Vol. 12, No. 1, p. 116, 2022.

[3] C. Sabanayagam, D. Xu, D. S. Ting, S. Nusinovici, R. Banu, H. Hamzah, and T. Y. Wong, "A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations", *The Lancet Digital Health*, Vol. 2, No.6, pp. e295-e302, 2020.

[4] S. W. Cheo, Q. J. Low, T. H. Lim, W. W. Mak, C. A. K. Yip, and K. W. Wong, "A practical approach to chronic kidney disease in primary care", *Malaysian Family Physician: the Official Journal of the Academy of Family Physicians of Malaysia*, Vol. 17, No. 1, p. 10, 2022.

[5] A. Nithya, A. Appathurai, N. Venkatadri, D. R. Ramji, and C. A. Palagan, "Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images", *Measurement*, Vol. 149, p. 106952, 2020.

[6] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng, and B. Chen, "A machine learning methodology for diagnosing chronic kidney disease", *IEEE Access*, Vol. 8, pp. 20991-21002, 2019.

[7] A. Khamparia, G. Saini, B. Pandey, S. Tiwari, D. Gupta, and A. Khanna, "KDSAE: Chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network", *Multimedia Tools and Applications*, Vol. 79, pp. 35425-35440, 2020.

[8] E. M. Senan, M. H. A. Adhaileh, F. W. Alsaade, T. H. Aldhyani, A. A. Alqarni, N. Alsharif, and M. Y. Alzahrani, "Diagnosis of chronic kidney disease using practical classification algorithms and recursive feature elimination techniques", *Journal of Healthcare Engineering*, 2021.

[9] K. Kalaiselvi and S. B. V. J. Sara, "A hybrid filter wrapper embedded-based feature selection for selecting important attributes and predicting chronic kidney disease", In: *Proc. of International Conference on Computing, Communication, Electrical and Biomedical Systems. Cham: Springer International Publishing*, pp. 137-153, 2022.

[10] B. M. Henry and G. Lippi, "Chronic kidney disease is associated with severe coronavirus disease (COVID-19) infection", *International Urology and Nephrology*, Vol. 52, pp. 1193-1194, 2019.

[11] M. Elhoseny, K. Shankar, and J. Uthayakumar, "Intelligent diagnostic prediction and classification system for chronic kidney disease", *Scientific Reports*, Vol. 9, No. 1, p. 9583, 2019.

[12] E. H. A. Rady and A. S. Anwar, "Prediction of kidney disease stages using data mining algorithms", *Informatics in Medicine Unlocked*, Vol. 15, p. 100178, 2019.

[13] A. Ogunleye and Q. G. Wang, "XGBoost model for chronic kidney disease diagnosis",

IEEE/ACM Transactions on Computational Biology and Bioinformatics, Vol. 17, No. 6, pp. 2131-2140, 2019.

[14] A. S. Shanthakumari and R. Jayakarthik, "Utilizing support vector machines for predictive analytics in chronic kidney diseases", *Materials Today: Proceedings*, Vol. 81, pp. 951-956, 2021.

[15] S. Srivastava, R. K. Yadav, V. Narayan, and P. K. Mall, "An Ensemble Learning Approach for Chronic Kidney Disease Classification", *Journal of Pharmaceutical Negative Results*, pp. 2401-2409, 2022.

[16] E. M. Senan, M. H. A. Adhaileh, F. W. Alsaade, T. H. Aldhyani, A. A. Alqarni, N. Alsharif, and M. Y. Alzahrani, "Diagnosis of chronic kidney disease using practical classification algorithms and recursive feature elimination techniques", *Journal of Healthcare Engineering*, 2021.

[17] M. Hosseinzadeh, J. Koohpayehzadeh, A. O. Bali, P. Asghari, A. Souri, A. Mazaherinezhad, and R. Rawassizadeh, "A diagnostic prediction model for chronic kidney disease in the Internet of Things platform", *Multimedia Tools and Applications*, Vol. 80, pp. 16933-16950, 2021.

[18] S. M. M. Elkholy, A. Rezk, and A. A. E. F. Saleh, "Early prediction of chronic kidney disease using deep belief network", *IEEE Access*, Vol. 9, pp. 135542-135549, 2021.

[19] M. Elhoseny, K. Shankar, and J. Uthayakumar, "Intelligent diagnostic prediction and classification system for chronic kidney disease", *Scientific Reports*, Vol. 9, No.1, p. 9583, 2019.

[20] A. Abdelaziz, A. S. Salama, A. M. Riad, and A. N. Mahmoud, "A machine learning model for predicting chronic kidney disease-based Internet of things and cloud computing in smart cities", *Security in Intelligent Cities: Models, Applications, and Challenges*, pp. 93-114, 2019.

[21] S. Yang, F. Lee, R. Miao, J. Cai, L. Chen, W. Yao, and Q. Chen, "RS-CapsNet: an advanced capsule network", *IEEE Access*, Vol. 8, pp. 85007-85018, 2020.

[22] M. Sabahno and F. Safara, "ISHO: improved spotted hyena optimization algorithm for phishing website detection", *Multimedia Tools and Applications*, Vol. 81, No. 24, pp. 34677-34696, 2022.

[23] F. Jiang, X. Zhi, X. Ding, W. Tong, and Y. Bian, "DLU-Net for pancreatic cancer segmentation", In: *Proc. of 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1024-1028, 2020.

[24] G. Mary and N. Suganthi, "Detection of Parkinson's disease with Multiple Feature Extraction Models and Darknet CNN Classification", *Computer Systems Science & Engineering*, Vol. 43, No. 1, 2022.