Research Article

Diabetes Detection using Convolutional Neural Network through Feature Sequencing

Vivek Vaidya[a] , and Dr L K Vishwamitra[b]
[a]
Ph.D. Scholar ,Deptt of C.S.E,,Oriental University, Indore(M.P.) India.
[b]Prof. ,Deptt of C.S.E. , Oriental University, Indore(M.P.) India. lkvishwamitra@gmail.com

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: In this work, we design a multi-scale convolutional neural networks (MCNNs) model for the previous detection of Diabetes mellitus. The detection technique is based on the proposed model's training by analyzing the data taken from non-diabetic and diabetic patients from PIMA, an Indian database. To achieve a high level of accurate training data for diabetes detection, we finally detect the parameters like diastolic blood pressure, body mass index (BMX), triceps skin thickness, number of pregnancies, inheritance and age factor. Based on CNN acting at different resolutions, the proposed architecture avoids the traditional step of manual extraction of features by extracting features and classifying them at one time within the same network of data neurons. The proposed approach provides better classification results than the usual methods for safer diagnosis of Diabetes.

Keywords: I.D.F., CNN, PIMA Database, ReLU**.**

1. Introduction

Diabetes is a disease that creates many other complexities in the body like cardiovascular and kidney problems, retinopathy, damage to our nervous system, among others, which can even lead to death [1]. This disease not only deadly but also cover a significant portion of society. The I.D.F. (2017) [2] places India as the number one country with the highest number of diabetic citizens. However, worldwide, the number of people with Diabetes is estimated at 387 million people whose ages range between 20 and 79. Of these, 46% are still not diagnosed [3]. These facts insist us to have a system or technique to detect Diabetes at its early stage. Recently, doctors use three methods to find the symptoms of these diseases areas: a) Fasting blood glucose testb) Oral glucose tolerance testc) Assessment of glycosylated haemoglobin [4].

There are two subtypes of Diabetes: type-1 and type-2 Diabetes. Other manifestations of Diabetes are gestational Diabetes, which is defined as early glucose intolerance or first detected during pregnancy or genetic defects of the cells of the pancreas, genetic defects of insulin function, other specific types of diseases such as pancreatic exocrine, endocrine disorders, or Abnormal infection forms of diabetes or diabetes mellitus associated with genetic syndrome [5].

Two aspects that characterize type 2 diabetes are; insulin resistance and a relational deficiency of insulin exudation. Its appearance is slow: it can evolve with degree hyperglycemia sufficient to cause organic and functional damage in patients many tissues but without clinical symptoms and therefore without a diagnosis for several years [5]. This form of Diabetes is most commonly established in adults and very mostly overweight 4 will be based on the diabetes database Indian PIMA from the U.C.I. Machine Learning Database [7].

In reviewing the contributions that have addressed the development of intelligent systems applied to the detection of Diabetes in the last 20 years, it was found that, although some detection systems based on conventional techniques have been designed [8] and others from the perspective of intelligent systems [9], [10], [11] these have not been explored in their entirety.

This paper develops a vision-based method for accurately diagnosing Diabetes. Our proposed model is based on neural networks that perform both classification and automated feature extraction. A multidimensional approach is proposed using multiple deep networks of different sizes (input size), and its outputs are integrated.

Section 1 describes the introduction and survey of the documents, followed by Section 2 containing the proposed method consisting of the database and the Conventional Neural Network (CNN). At the same time, the simulation and results are presented in Section 3. Work.

2. Proposed Methodology

The methodology used is composed of the following steps: 1) Definition of the databases; 2) Structure of the Convolutional Neural Networks classifier; and 3) CNN Based on Feature Sequencing

A. PIMA Database

The tests of the proposed method are based on Bima Indians diabetes data [7]. Database selected from U.C.I. 768 Bima is a repository study for Indian women (500 non-diabetic 268 diabetics). The same women who stopped immigrating to Arizona (U.S.A.) follow a Western lifestyle, developing Diabetes in almost 50% of cases. Diagnosis is a different binary value "class" that allows the patient to identify whether they have symptoms of Diabetes following the functions of the World Health Organization.

The eight clinical descriptors are:

1. Npreg: number of pregnancies.
2. Glu: concentration of plasma glucose.
3. B.P.: diastolic blood pressure (mmHg).

4. SKIN: triceps skinfold thickness, (mm).
5. Insulin: insulin dose, (mu U / ml).
6. B.M.I.: body mass index, (weight in kg / (height in m 2).
7. D.P.F.: Diabetes pedigree function (heredity).
8. Age: age (Year).

B. Analysis of Database

Table 1 contains information on the parameters taken into consideration.

Table 1: Descriptor information in the database [7]

Attribute Name	**Min/Max**	**Standard deviation**	**Segregated**
Npreg	0 / 17	3.37	17
Glu	0 / 199	31.973	136
BP	0 / 122	19.356	47
Skin	0 / 99	15.952	51
Insu	0 / 846	115.244	186
Bmi	0 / 67.1	7.884	248
Ped	0.078/2.42	0.331	517
Age	21 / 81	11.76	52

C. Convolutional Neural Networks

Conventional neural networks or CNNs are a particular type of neural networks by its initials recommended for processing data by topology in the form of a mesh or grid. The data type most commonly used with these types of networks (meshes of x and y pixels), however, is also used for time series (data is a temporary dimension in one dimension with an additional dimension), as magnetic resonance scanners or videos in three dimensions (two dimensions associated with images and one dimension associated with the temporary development of video) [12].

CNN's have had great success with many missions. Recently, human vision has improved due to a profoundly changeable neural network based on image recognition [13].

History and Development

Evolutionary neural networks have played a significant role in the history and evolution of artificial neural networks. They are a clear example of the use of brain biology and physiology (CNNs have similarities to human vision) to create artificial mechanisms within the realm of machine learning, and especially deep learning.

They were one of the first models of neural networks with good results and performance, which were already used to develop commercial applications at the end of the last century. For example, in 1990, an AT&T research team developed an application to read invoices using modified neural networks. By the end of the 1990s, about 10 per cent of U.S. invoices were read by this method [14].

Many years later, Microsoft developed models for recognizing people's signatures. In 2012, when Khrushchevsky won the image collection competition ImageNet [15], one of the most outstanding and most recent achievements in the use of convertible neural networks was where you had to categorize numerous images into about a thousand different classes.

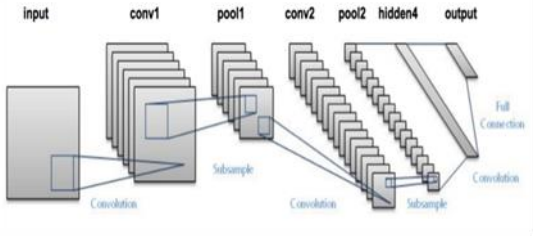

Figure 1: Typical architecture of a deep convolutional neural network [15]

The Operation of Convolution

In its most general form, a convolution is an operation applied to two functions with real numbers as arguments. The following mathematical expression defines the operation of convolution:

$$s(t) = \int x(a)w(t-a)da \quad (1)$$

Commonly the convolution operation is symbolized by:

$$s(t) = (x * w)(t) \quad (2)$$

Using the terminology associated with convolutional neural networks, the first term (in this case (x)) of the convolution operation is often referred to as input, while the second argument (in our case w) it's called the kernel. The output or result of the operation of convolution is usually called a feature map.

When working with a computer, discrete data will be available so that what used to be an integral function of logical functions will have to become a sum of "discrete" functions, also continuous, of the following shape:

$S(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t-a)$ (3)

In deep learning applications, the input is usually a vector of several dimensions (tensor). The kernel is often a multidimensional vector of parameters modified by the learning algorithm. For example, if it is used as input data, an image I, the most frequent is that a two-dimensional kernel is used, which in this case we will denote as K:

$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(i-m, j-n)K(m,n)$ (4)

In practice, the discrete convolution can be considered as a multiplication by a matrix:

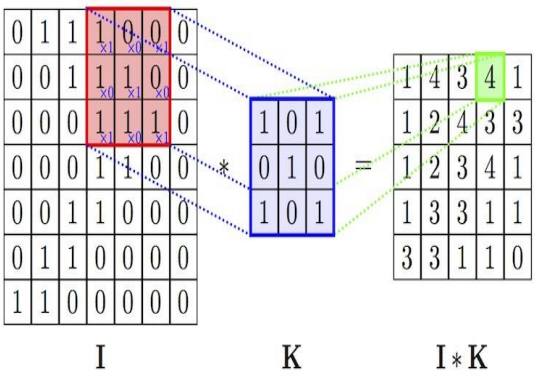

Figure 2: Operation performed by a two-dimensional convolutional layer of a convolutional neural network – CNN

Taking Figure 2 as a reference, we observe how the two-dimensional input I remains fixed, while the K kernel, which is also a matrix of two dimensions, moves along the entrance I, performing operations of multiplication element by element and adding the total, obtaining the result of the convolution (I×K).

Pooling

Three levels typically form a convolutional neural network. In the first layer of the network, the convolution operations take place on the input data. In the second stage, each feature extracted by the convolution operation is passed to an activation function, the most used being the ReLU (Rectified Linear Unit) activation function. This stage is sometimes called the detector stage [16].

In the last stage, what is known as the pooling function is performed, which replaces the output or output of the network with a statistical summary made by zones of the previous layer of the neural network. It is much easier to understand with Figure 3:

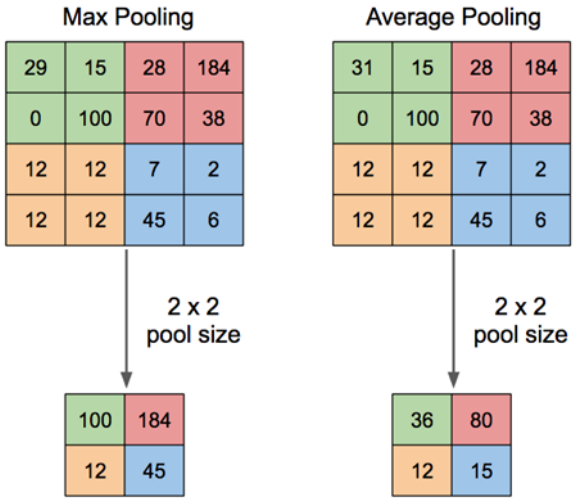

Figure 3: Pooling function, where on the left side a max-pooling is observed, and on the right side there is an average-pooling [16]

The figure shows how the pooling operation is applied in 2×2 areas; there are two types of possible operations:

Max-pooling: In this operation, the highest value is chosen within all possible, so the reduction of characteristics or values is a factor of four.

Average-pooling: In this operation, the arithmetic average of the values within the region to be applied is chosen, also obtaining a data reduction factor of 4.

D) **CNN Based on Feature Sequencing:**

Framework. This paper uses the CNN-based model to apply low-dimensional source features directly to the sample input. A feature rank level is added to adjust the order of features automatically. This approach can save variable time, use CNN, learn derivative features useful for classification results, and reduce human experience interference

with the model. The goal of minimizing interference is to allow as much learning as possible of CNN transaction characteristics and features mode.

The model's overall structure is divided into two parts: the sample's training area and the diabetes diagnosis area. The training area of the model is divided into two parts: feature sequence level and CNN. An enhanced feature sequence is used to improve the order of symptoms of the disease. First, place the data in the feature sequence layer, the sample effect is tested by tutoring the CNN model, and the feedback changes the sequence effect of the feature sequence.

At the time of update, the standard feature queue time can find the best order mode. When entering the real-time data model, the data features are sorted by aspect sequence. Then the training model is determined between the multidimensional features so that the multidimensional properties can be almost configured. If transaction data is placed in the algorithms of different models in the form of one-dimensional variables, the configuration and the combination of different properties do not affect the record's physical meaning. But different permutations and combinations can affect the results of the model.

Algorithm to find the best permutation for input feature:

Input: There is weight matrix W, the row of weight matrix represents the initial stage of input set S are, $W_1, W_2, \dots \dots W_n$. The accuracy is represented as Z. auxillary sequence $a_1, a_2, \dots . a_n\left(a_j \; is \; the \; nummber \; of \; rows \; below \; W_j, satify \; the \; first \; condition \; 0 \le a_j \le n\right)$ and $p_1, p_2, \dots p_n . P_j \; control \; the \; direction \; of \; a_j changes$.

Output: The best permutation of the weight matrix

1. $a_j \leftarrow 0$
2. $p_j \leftarrow 1 (1 \le j \le n)$
3. $Z \leftarrow [\,]$
4. $Access$ matrix W
5. $j \leftarrow n . s \leftarrow 0 (\; s \; is \; the \; number \; of \; a_k \; stisfying \; k > j \; and \; a_k = k - 1)$
6. $q \leftarrow a_j + p_j;$
7. $if \; q < j$
8. Go to 19
9. End if
10. $if \; q = j$
11. Go to 14.
12. End if;
13. $W_{i-a_j+s} \leftrightarrow W_{j-q+s} \leftarrow q, go \; to \; 4$
14. $if \; j = 1, then$
15. Input S and W model to calculate accuracy
16. Z.append (accuracy)
17. If $len(Z) == n!$ Then
18. Return max (Z) and W that has max (Z)
19. End if.
20. Else
21. s=s+1
22. end if;
23. $p_j = -p_j, j = j - 1, go \; to \; 6.$

3. SIMULATION RESULTS

A. Evaluation Criteria

Data classification performance was evaluated by calculating true positives (T.P.s), true negatives (T.N.s), false positives (F.P.s) and false negatives (F.N.s), per cent sensitivity (S_e), specificity (S_p) and the classification rate (T.C.), their respective definitions are as follows:

- VP: diabetic classified diabetic.
- VN: non-diabetic classified non-diabetic.
- FP: non-diabetic classified diabetic.
- FN: diabetic classified as non-diabetic.

Sensitivity is the capacity to provide a positive outcome when the disease is present. It is calculated by:

$$S_e = \frac{VP}{VP+FN} \quad (5)$$

Specificity is the capacity to provide a negative outcome when the disease is absent. It is calculated by:

$$S_p = \frac{VN}{VN+FP} \quad (6)$$

Classification rate is the percentage of correctly classified examples. It is calculated by:

$$TC = \frac{VP+VN}{VN+VN+FP+FN} \quad (7)$$

B. Results and Interpretation

Table 2: Result for proposed work

Epoch	**Iteration (hh:mm:ss)**	**Time Elapsed**	**Mini-batch Accuracy**	**Mini-batch Loss**	**Base Learning Rate**
1	1	00:00:02	49.22%	0.6983	0.0100
17	50	00:00:07	95.31%	0.0798	0.0100
34	100	00:00:07	98.44%	0.0450	0.0100
50	150	00:00:08	98.44%	0.0254	0.0100
67	200	00:00:09	96.88%	0.0545	0.0100
84	250	00:00:09	98.44%	0.0382	0.0100
100	300	00:00:10	100.00%	0.0183	0.0100
117	350	00:00:10	97.66%	0.0478	0.0100
134	400	00:00:11	98.44%	0.0325	0.0100
150	450	00:00:11	100.00%	0.0140	0.0100

Table 3: Convolutional Neural Network Layers

'imageinput'	Image Input	4×2×1images with 'zerocenter' normalization
'conv'	Convolution	52×2×1 convolutions with stride [1 1] and padding [0 0 0 0]
'relu'	ReLU	ReLU
'fc'	Fully Connected	2 fully connected layer
'softmax'	Softmax	Softmax
'classoutput'	Classification Output	crossentropyex with classes '0' and '1'

Here 67 outputs out of 73 on the Multi-Scale Convulsive Neural Network will give the correct result, and the remaining 6 will provide the wrong output, so the exact number for non-diabetics is 91.8%.

Now 227 out of 234 cases of people with Diabetes give a correct result, and the remaining 7 give wrong results, so the accuracy of the person with Diabetes is 97%. So the exact number drawn for both diabetics and non-diabetics is 95.76%.

Error Rate= $\frac{\text{FN+FP}}{\text{TOTAL}} = \frac{7+6}{307}$ =4.2%

Miss Rate= $\frac{\text{FN}}{\text{TP+FN}} = \frac{7}{227+7}$ =3%

False Positive Rate = $\frac{\text{FP}}{\text{FP+TN}} = \frac{6}{6+67}$ =8.2%

Specificity= $\frac{\text{TN}}{\text{FP+TN}} = \frac{67}{6+67}$ =91.8%

True Positive Rate= $\frac{\text{TP}}{\text{TP+FN}} = \frac{227}{227+7}$ =97%

Accuracy= $\frac{TP+TN}{Total} = \frac{227+67}{307}$ =95.76%

This approach shows that the uniqueness of the system is very high, i.e. the system is well trained for negative data. Therefore, when a patient has Diabetes, our model will diagnose it very successfully.

Sensitivity of the system is fragile. The system incorrectly recognizes the positive data. Many people with Diabetes are identified as people with Diabetes. This can pose a significant risk to the patient's health.

Our model performance gave an average classification rate and a good individuality.

Table 4: Performance Table

Method	**Number of Attributes**	**Error rate**	**Specificity**	**Accuracy**
CNN (Proposed)	8	4.2%	91.8%	95.76 %

C. Comparison with Works of Literature

This approach presents a comparative analysis of the work done in this field with the results obtained from our approach and the BIMA database. Table 5 summarizes the comparisons with other works:

Table 5: Comparison Table with Literature Works

Name of The Method	**Type of Method**	**Number of Attributes**	**Accuracy**
FCBF+ SVM [17]	Filter	4	77.99 %

GR+RBF [18]	Filter	5	86.46 %
GR+ MLP [18]	Filter	5	78.21 %
CAFS+ MLP [19]	Wrapper	6	76.18 %
GA+SVM [20]	Wrapper	4	81.50 %
TS1+ MLP [21]	Wrapper	4	79.55 %
CNN (Proposed)	Multi-Scale Convolutional Neural Network	8	95.76 %

4. Conclusion

Proposed method that confirms its performance in testing and provided interesting improvements in-sample error, sensitivity and classification ratios while improving its structure. After comparing the results obtained by literary works, we have noticed that the results obtained are comparable or better than other results. To test these results, we consulted with people with Diabetes who confirmed that the three types they had chosen were the most appropriate for their diagnosis of Diabetes.

To develop a powerful application that can be used to diagnose Diabetes, we have implemented a variable selection system to eliminate powerless variables that discriminate between classes. Good education is essential. This method selected a variable (glucose: glucose, B.M.I.: mass, P.E.T.: heredity) as the most appropriate variables to create a better classification. We consulted with people with Diabetes to verify these results, confirming that all three types are highly appropriate for the diagnosis of Diabetes.

The results obtained after using this method are very promising. They are based on the work already done in this field, which confirms the seriousness of the proposed contribution to our problem's solution. We plan to confirm the interpretation of the multidimensional neural network classifier of the model in the future. We would like to generalize this common use of all types of diseases, which should be integrated into appropriate diagnostic support systems in future hospital or clinic office fees.

References

1. Beran, David, and John S. Yudkin. "Diabetes care in sub-Saharan Africa." The Lancet 368, no. 9548 (2006): 1689-1695.
2. Annual Report 2017- International Diabetes Federation. Online Available at: https://www.idf.org/component/attachments/attachments.html?id=1851&task=download
3. American Diabetes Association, 2015. 2. Classification and diagnosis of diabetes. Diabetes care, 38(Supplement 1), pp.S8-S16.
4. Beran, David, John S. Yudkin, and Maximilian De Courten. "Access to care for patients with insulin-requiring diabetes in developing countries: case studies of Mozambique and Zambia." Diabetes care 28, no. 9 (2005): 2136-2140.
5. American Diabetes Association, 2014. Standards of medical care in diabetes—2014. Diabetes care, 37(Supplement 1), pp.S14-S80.
6. Durairaj, M. and Kalaiselvi, G., 2015. Prediction of diabetes using soft computing techniques-A survey. International journal of scientific & technology research, 4(3), pp.190-192.
7. "Pima," A. Frank and A. Asuncion, 2010, Pima Indians Diabetes Data Set, UCI Machine Learning Repository, University of California, Irvine, School of Information and Computer Sciences, [http://archive.ics.uci.edu/ml].
8. Torres-Zapata, A. E., M. A. Aparicio-Trápala, J. L. Blé-Castillo, and C. A. Corzo-Sosa. "Glycemic and insulinic response of patients with type 2 diabetes to the consumption of pumpkin soup creole (CucúrbitaPepo L.) enriched with banana starch. RespuestaGlucémica e Insulínica de Pacientes con Diabetes tipo 2 al Consumo de sopa de Calabaza Criolla (CucúrbitaPepo L.) Enriquecida con Almidón de Banano." Inf. Tecnol 23 (2012): 71-86.
9. Herrero, Pau, Jorge Bondia, OloruntobaAdewuyi, Peter Pesl, Mohamed El-Sharkawy, Monika Reddy, Chris Toumazou, Nick Oliver, and Pantelis Georgiou. "Enhancing automatic closed-loop glucose control in type 1 diabetes with an adaptive meal bolus calculator–in silico evaluation under intra-day variability." Computer methods and programs in biomedicine 146 (2017): 125-131.
10. Kavakiotis, Ioannis, Olga Tsave, Athanasios Salifoglou, Nicos Maglaveras, Ioannis Vlahavas, and IoannaChouvarda. "Machine learning and data mining methods in diabetes research." Computational and structural biotechnology journal 15 (2017): 104-116.
11. Gargeya, Rishab, and Theodore Leng. "Automated identification of diabetic retinopathy using deep learning." Ophthalmology 124, no. 7 (2017): 962-969.
12. Shen, Wei, Mu Zhou, Feng Yang, Caiyun Yang, and JieTian. "Multi-scale convolutional neural networks for lung nodule classification." In International Conference on Information Processing in Medical Imaging, pp. 588-599. Springer, Cham, 2015.

13. Cai, Zhaowei, Quanfu Fan, Rogerio S. Feris, and NunoVasconcelos. "A unified multi-scale deep convolutional neural network for fast object detection." In European Conference on Computer Vision, pp. 354-370. Springer, Cham, 2016.
14. Yang, Wei, Yingyin Chen, Yunbi Liu, Liming Zhong, Genggeng Qin, Zhentai Lu, Qianjin Feng, and Wufan Chen. "Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain." Medical image analysis 35 (2017): 421-433.
15. Krizhevsky, Alex, IlyaSutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In Advances in neural information processing systems, pp. 1097-1105. 2012.
16. Zeiler, Matthew D., and Rob Fergus. "Stochastic pooling for regularization of deep convolutional neural networks." arXiv preprint arXiv:1301.3557 (2013).
17. Balakrishnan, S. and Narayanaswamy, R., 2009. Feature selection using FCBF in type ii diabetes databases. International Journal of the Computer, the Internet and the Management, 17(1), pp.50-8.
18. Karegowda, A.G., Manjunath, A.S. and Jayaram, M.A., 2010. Comparative study of attribute selection using gain ratio and correlation based feature selection. International Journal of Information Technology and Knowledge Management, 2(2), pp.271-277.
19. Kabir, M.M., Islam, M.M. and Murase, K., 2010. A new wrapper feature selection approach using neural network. Neurocomputing, 73(16-18), pp.3273-3283.
20. Huang, C.L. and Wang, C.J., 2006. A GA-based feature selection and parameters optimizationfor support vector machines. Expert Systems with applications, 31(2), pp.231-240.
21. Wang, Y., Li, L., Ni, J. and Huang, S., 2009. Feature selection using tabu search with long-term memories and probabilistic neural networks. Pattern Recognition Letters, 30(7), pp.661-670.