

Article

Deep Learning Methods for Classification of Certain Abnormalities in Echocardiography

Imayanmosha Wahlang ¹, Arnab Kumar Maji ¹, Goutam Saha ¹, Prasun Chakrabarti ², Michal Jasinski ^{3,*}, Zbigniew Leonowicz ³ and Elzbieta Jasinska ⁴

¹ Department of Information Technology, North-Eastern Hill University, Shillong 793022, Meghalaya, India; imayanwahlang@gmail.com (I.W.); arnab.maji@gmail.com (A.K.M.); dr.goutamsaha@gmail.com (G.S.)

² Techno India NJR Institute of Technology, Udaipur 313003, Rajasthan, India; drprasun.cse@gmail.com

³ Department of Electrical Engineering Fundamentals, Faculty of Electrical Engineering, Wroclaw University of Science and Technology, 50-370 Wroclaw, Poland; zbigniew.leonowicz@pwr.edu.pl

⁴ Faculty of Law, Administration and Economics, University of Wroclaw, 50-145 Wroclaw, Poland; elzbieta.jasinska@uwr.edu.pl

* Correspondence: michal.jasinski@pwr.edu.pl

Abstract: This article experiments with deep learning methodologies in echocardiogram (echo), a promising and vigorously researched technique in the preponderance field. This paper involves two different kinds of classification in the echo. Firstly, classification into normal (absence of abnormalities) or abnormal (presence of abnormalities) has been done, using 2D echo images, 3D Doppler images, and videographic images. Secondly, based on different types of regurgitation, namely, Mitral Regurgitation (MR), Aortic Regurgitation (AR), Tricuspid Regurgitation (TR), and a combination of the three types of regurgitation are classified using videographic echo images. Two deep-learning methodologies are used for these purposes, a Recurrent Neural Network (RNN) based methodology (Long Short Term Memory (LSTM)) and an Autoencoder based methodology (Variational AutoEncoder (VAE)). The use of videographic images distinguished this work from the existing work using SVM (Support Vector Machine) and also application of deep-learning methodologies is the first of many in this particular field. It was found that deep-learning methodologies perform better than SVM methodology in normal or abnormal classification. Overall, VAE performs better in 2D and 3D Doppler images (static images) while LSTM performs better in the case of videographic images.

Keywords: abnormalities; Convolutional Neural Network (CNN); echocardiogram; Long Short Term Memory (LSTM); regurgitation; Variational AutoEncoder (VAE)



Citation: Wahlang, I.; Maji, A.K.; Saha, G.; Chakrabarti, P.; Jasinski, M.; Leonowicz, Z.; Jasinska, E. Deep Learning Methods for Classification of Certain Abnormalities in Echocardiography. *Electronics* **2021**, *10*, 495. <http://doi.org/10.3390/electronics10040495>

Academic Editors: Giovanni Dimauro, Vitoantonio Bevilacqua and Leandro Pecchia

Received: 29 January 2021

Accepted: 17 February 2021

Published: 20 February 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the advances in the field of biomedical imaging, digital images play a vital role in the early detection of abnormalities or diseases in the human body for any systems. Many intricate systems exist in the human body, namely the nervous system, cardiac system, endocrine system, etc that are important for survival. Out of these, the cardiac system is considered to be one of the most delicate systems. Cardiology is viewed as a complex subject of practice due to less exposure to the intricacies of relevant technologies. Medical imaging has become a tool for diagnosis purposes and provides information about the anatomic structures with the assistance of computers through imaging modalities like Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Angiogram, Electrocardiograph (ECG) and others [1].

Amongst these, echocardiogram (echo) is considered and perhaps the most frequently used tool in the field of the cardiac system. It is used mainly due to its ability for early diagnosis and management of heart diseases. It is a simple, non-invasive, and inexpensive technique that can precisely show the pressure gradient of heart lesions. Since it uses sound waves instead of radiation, echo is considered to be safe [2]. Echo uses standard

two-dimensional (2D), three-dimensional (3D), and Doppler ultrasound to create images of the heart [3]. The echo plays a crucial role in the diagnosis of cardiac diseases. The images used in this work are captured when the patient is in the left lateral decubitus (when the patient is lying on the left side) [4]. The transducer used is transthoracic (without insertion of transducer inside the esophagus) [5,6]. The different parts of the heart are Left Ventricle (LV), Right Ventricle (RV), Left Atrium (LA), and Right Atrium (RA). Some different views/planes involved in echo are Parasternal long-axis view (PLAX), RV inflow view (RVIT), Parasternal short-axis view (PSAX), Apical view, Subcostal view etc. [7]. In PLAX view the transducer is placed towards the right shoulder, whereas the PSAX view is obtained by rotating the transducer by 90° in the clockwise direction from the PLAX view [8]. In the case of RVIT, the sound beam of the transducer is pointed towards the right hip [8]. The apical view is quite similar to the PLAX view. However, the only difference is the image is taken from the apex portion of the heart [8]. The subcostal view can be obtained by placing the transducer towards the left, and the sound beam is projected slightly anteriorly [8]. Echo has become a prominent choice used in the examination of valvular heart diseases (regurgitation). Regurgitation [9] is one of the most common valvular diseases that is characterized by the prominent use of tobacco and alcohol. It can be detected from abnormal flow patterns using color imaging (color flow mapping). The main types of regurgitation are:

1. Mitral Regurgitation (MR): It is the most common valvular involvement in children and rheumatic heart diseases. In color flow, LA size is increased, and the MR jet can be seen [10].
2. Aortic regurgitation (AR): AR is known to be less frequent than MR. But most patients having AR have associated mitral valve disease. It results from distorted aortic leaflets and for which careful analysis of the aortic valve is a must [9]. Color flow and Doppler give an estimate of the severity of AR.
3. Tricuspid Regurgitation (TR): It is often common in people who smoke [11]. It can also be seen in 20% of rheumatic heart disease patients. Using Doppler and color flow, TR can be seen, and depending on the TR jet, the severity of TR can be revealed. The tricuspid valve is similar to that of the mitral valve with more variability [9].

The above mentioned three types of regurgitation are acquired heart disease (cause during one's lifetime). Not all kinds of regurgitation are known to be acquired, as some can be congenital (presence since birth). In our case, it was observed that all cases of patient data are acquired. Figure 1 shows a Doppler echo having AR, MR, and TR abnormalities respectively. This valvular regurgitation plays a significant role and represents an important cause in mortality and morbidity [9]. The echo plays an essential role in regurgitation assessment and using Doppler echo the presence of the types of regurgitation can be distinguished more distinctively. But this has to be done by a cardiologist by precisely locating and assessing the visualization in the form of video. To detect the presence or absence of any abnormalities, extraction of an image or images from a videographic echo is a necessity. From the visualization of echo, a cardiologist can predict the functions of valves and defective parts, if any. However, It requires trained cardiologists to interpret accurate findings and give reports. Often cardiologists take the help of cauterization [12], which is a surgically invasive and expensive procedure. Usage of automated methods will help in the accurate diagnosis of any heart abnormalities and also reduce the necessity of invasive procedures. There are no automated facilities that can detect the presence of abnormalities or any disease in the heart. Thus, finding a way to treat such abnormalities using automated algorithms is needed. An attempt using machine learning algorithms have been made in the past in which SVM has provided better result using static images. But using videographic images was never explored. For this purpose, work using videography has been introduced to reduce the work of a cardiologist and provide an efficient and effective result as this will help in the early detection and diagnosis of heart diseases.

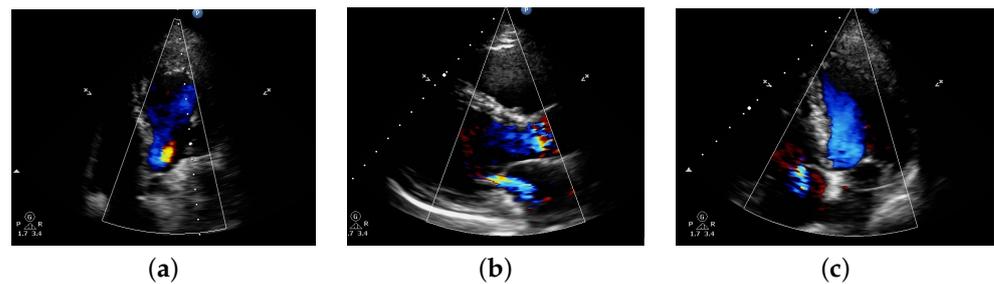


Figure 1. Diagram showing Doppler echo from dataset collected of patient having (a) Aortic Regurgitation (AR), (b) Mitral Regurgitation (MR), and (c) Tricuspid Regurgitation (TR) abnormalities respectively.

Usage of deep learning and machine learning techniques help us to a greater extent to handle such fine details. This work aims to classify images into normal and abnormal and the different type of regurgitation. It has been observed that the performance of small architecture is quite similar with the performance of more complex architecture. We wanted to examine the clinical capability of our method in classifying the different types of images. Firstly, an effort has been made to classify the echo images into two classes (i.e., abnormal and normal). Then videographic echo images have been considered for further classifications, based on regurgitation present, using two types of deep learning-based models (i.e., LSTM and VAE-CNN). Firstly, RNN based model using LSTM has been chosen due to its capability in recalling the time and predicting the next image or frame. Another method is using Variational Auto Encoder (VAE) with Convolutional Neural Network (CNN). CNN is used for extracting features and for space reduction. A comparison with well-known SVM methodology is also performed using the static images of the echo.

The main contributions of the paper are as follows:

1. Work based on videographic images has been proposed as an initiative to find out its usefulness in the diagnosis of different types of abnormalities. Videographic images are used for classification into six types of regurgitation and two-class (normal or abnormal) classification.
2. Work on 2D images for classification into normal or abnormal in PLAX view [13] has been done and compared to an existing method, i.e., SVM [14,15].
3. Using color Doppler 3D images, classification into normal or abnormal was done.
4. Using RNN and CNN based VAE deep learning methodologies including an existing technique SVM, used for 2D classification, was done.
5. Classification is performed, using the images captured in the Radiology laboratory and validated with the help of a cardiologist.

Works related to regurgitation classification are described in Section 2. In Section 3, the flowchart of different methodologies used, are explained in brief. Section 4 provides the experimental result and Section 5 consists of the conclusion and future work.

2. Related Works

Work based on the cardiac system has become one of the most popular and aspiring field for many researchers. This is because it is one of the most important system of the human body and is the leading cause for morbidity and mortality in patients with kidney disease in the United States [16]. It is one of the main organs for blood supply and can be called as a manufacturer of blood circulation and thus plays a vital role. During echo visualization, any abnormal inflow or outflow of blood can be a sign of abnormalities or diseases in the heart. For this reason, works related to heart abnormalities have been taken in this paper and are discussed further. Many of the related works are not based on the classification of heart abnormalities but are included as they deal with the classification in the heart-related field.

Work related to cardiac classification can be found in Allan et al. [14], where the classification of Mitral Regurgitation (MR) was carried out using SVM as a classification method with an accuracy of 82% for moderate or severe MR. The apical view was taken for

this purpose using a 2D echo with 6993 studies obtained from the Clinical Medical Research Ethics Board of Vancouver Coastal Health [14]. Balaji et al. have done works on the classification of different views of echo wherein [17] parasternal short axis (PSAX), parasternal long axis (PLAX), apical two-chamber (A2C), and apical four-chamber (A4C) was classified using the histogram and statistical features with 87.5% accuracy and in [18], parasternal long axis (PLAX), apical two-chamber (A2C) and apical four-chamber (A4C) was classified using Connected Component Labelling with 94.56% accuracy. Nandagopalan [19], has also worked for view classification where parasternal short axis (PSAX), parasternal long axis (PLAX), apical two-chamber (A2C), and apical four-chamber (A4C) was classified using a proposed method with 96% accuracy. Pinjari [20] used Proximal Isovelocity Surface Area (PISA) method for the classification of mild Mitral Regurgitation (MR), moderate Aortic Regurgitation (AR) and severe Aortic Regurgitation (AR). It was done using color Doppler images where images used were MR and AR. The images were first converted into YCbCr space, and filtering techniques like Wiener and Gaussian filters were applied. For the same, Segmentation was done using Fuzzy C Means. Another work can be seen in [21], where heart valve disease for AR is assessed. It is assessed using a gradient, Aortic Stenosis (AS) grading, peak velocity, velocity ratio, Aortic Valve Area (AVA), Indexed AVA, and mean gradient. The type of AS is known based on the ratio obtained. Also, Strunic et al. worked on the classification of murmurs using ANN as a classification technique using heart sound [22]. Many papers have considered Left Ventricle (LV) segmentation as an important aspect in finding abnormalities considering LV as the largest part of the heart where the flow of blood can be witnessed [23,24]. A review work was done on machine learning for heart disease prediction in [25], and work on the classification of heart diseases based on the counts of heart beat could be seen in [26].

Along with these methods, other state-of-art emerging techniques are deep learning like Convolutional Neural Network (CNN), Autoencoders, Recurrent Neural Network (RNN) in different fields of biomedical imaging, and computer vision. Deep learning has several families, including fully connected networks like autoencoders, convolutional neural networks like AlexNet, and LeNet, recurrent neural networks like LSTM, and deep belief networks. Using Deep Learning architecture like CNN has an advantage over other Deep learning methodologies, where features are extracted during the process. These architectures have shown excellent performance in many fields and even gained popularity in the field of segmentation of images.

Works related to normal or abnormal heart images have not been done previously. Such work is necessary to help physicians for identification of the presence or absence of any abnormalities. This work has been taken in this paper with a hope that automated methodologies can reduce human exertion and applicable as a tool in places abstain by an expert or ease the process of diagnosis.

In this paper, work based on classification has been taken as an initial step for the prediction of a specific region of interest. Two types of classification have been carried out, namely, classification into normal and abnormal images and classification into different types of regurgitation.

3. Classification of Heart Abnormalities Using Different Architectures

Classification plays an important role in the prediction of an area or region containing abnormalities for the diagnosis of any disease. It classifies input into different classes. In this work, we have used Long Short Term Memory (LSTM), Variational Autoencoder + Convolutional Neural Network (VAE-CNN) along with SVM are used for classification. The 2D static echo images and 3D static doppler images are classified into two classes namely normal or abnormal. Videographic echo images are also classified into two classes (normal or abnormal) and six-classes of regurgitation using the same methodologies.

3.1. Data Acquisition

The raw data were obtained from a Cardiac Clinic namely Hope Clinic located in Shillong, India using echo as a tool under the supervision of specialist in the relevant field.

A sample image used in the work is shown in Figure 1. The data obtained are in 2D jpeg images, 3D bitmap images, colored and 2D videographic images in Audio Video Interleave (AVI) format. A total of 120 patient data with abnormality/abnormalities cases and a few normal cases were collected. The different types of abnormalities are MR, AR, TR, and a few having mixture of these. All the data are validated with the help of a cardiologist.

3.2. Image Preprocessing and Data Augmentation

An overall flowchart depicting the working methodologies is shown in Figure 2. Our scheme starts with taking an input image (frame in case of video), which is then cropped and converted into gray scale for 2D classification of images and videos. This conversion is important as grayscale images are more detailed and give a better representation of the image. It is then passed for filtering using the Gaussian filter as in [20]. It was done to remove noise, and unwanted data and Gaussian filter give a comparatively better result. Few images having a mixture of two or more abnormalities were augmented so as to obtain 10 number of patient data for experimental purpose in the case of video classification of 6 classes. Augmentation using cropping has been done.

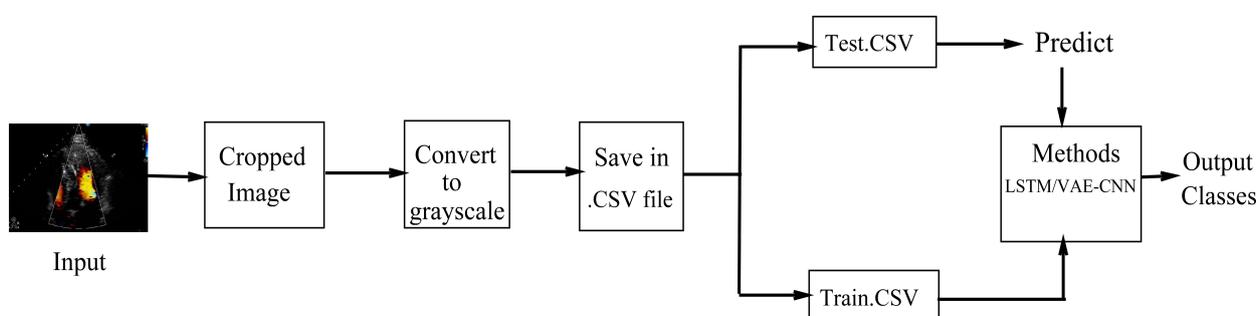


Figure 2. An overall flow chart showing the working methodologies used in our scheme.

3.3. Classification Using LSTM, VAE-CNN and SVM Methodologies

The images after preprocessing are then saved into two Comma-Separated Values (CSV) files for the testing and training phase. The training CSV file is used for the training and validation phase consisting of a labelled dataset. The images are then processed using different methodologies (LSTM/VAE-CNN/SVM) for validation purposes. The testing CSV file consists of unlabelled data. After which the test images will be predicted, and the output obtained is class 0 or 1 in the case of two class classification and class 0, 1, 2, 3, 4, 5 in the case of six classes for regurgitation classification.

Steps Involved in Classification of Video

The steps involved in Videographic images are as follows:

1. Extract each frame and operate on each of them. This is known as spatiotemporal deep learning.
2. Each frame is assigned a class in the training and validation phase (labelled frames).
3. Frames are cropped and resized into 224×224 . The size was chosen randomly based on the previous network, like AlexNet.
4. The training set is then passed into the network (LSTM and VAE-CNN) for classification.
5. Output classes 0, 1, 2, 3, 4, 5 in the case of six-class classification and 0 and 1 in the case of two-class classification were obtained.
6. Testing was done on the remaining unlabelled frames of each video.
7. Steps 3 to 5 are repeated.

3.3.1. Long Short Term Memory (LSTM)

LSTM is an RNN based model. LSTM methods are used in speech and Natural Language Processing (NLP). Since 1997, when Hochreiter introduces LSTM, it has become

prevalent in the field of text classification [27]. But this RNN technique has also been found suitable for videos by researchers and can help in predicting the next frame in a video and are applied in many fields of videography [28]. For this reason, this method has been used in our case.

This paper aims to use RNN using LSTM as a variant as it is a better version of RNN [28]. RNN has many variants, including LSTM, GRU, and other modified versions. Here, taking video as input is challenging compared to images as videos are a collection of frames.

LSTM is designed to overcome long term dependencies and to solve the vanishing gradient problem [29]. LSTM improves gradient flow and is most suitable when time is taken as a factor. In a way, LSTM is similar to ResNet (Residual networks) [29].

In this paper, the LSTM model is used without any change in architecture. It has input gate units and output gate units, and the resulting units which are complicated are called memory cell [27]. It also consists of forget gate, memory cell inputs, memory cell output. Gates are used for the memorizing process [30]. A diagram showing the working components of LSTM is shown in Figure 3. The elements of LSTM can be calculated as:

$$F_t = \text{Sigmoid}[W_f(H_{t-1}, X_t)] \tag{1}$$

$$I_t = \text{Sigmoid}[W_i(H_{t-1}, X_t)] \tag{2}$$

$$G_t = \text{Sigmoid}[W_g(H_{t-1}, X_t)] \tag{3}$$

$$O_t = \text{Sigmoid}[W_o(H_{t-1}, X_t)] \tag{4}$$

$$C_t = F_t \cdot C_{t-1} + I_t \cdot G_t \tag{5}$$

$$H_t = O_t \cdot \tanh(C_t) \tag{6}$$

where F_t is the forget gate, I_t is the input state, G_t is the cell state, W is the weight, H is the output, X is the input, O_t is the output of the sigmoid gate, and C_t is the cell state. For our purpose, the input is passed to a convolutional layer for feature extraction which is then passed to LSTM architecture, and the output is classified into 0 or 1 (normal or abnormal) class for 2D classification, Doppler classification, and two-class video classification using Sigmoid classifier, and class 0 to 5 for six-class classification of regurgitation for videographic images using softmax classifier.

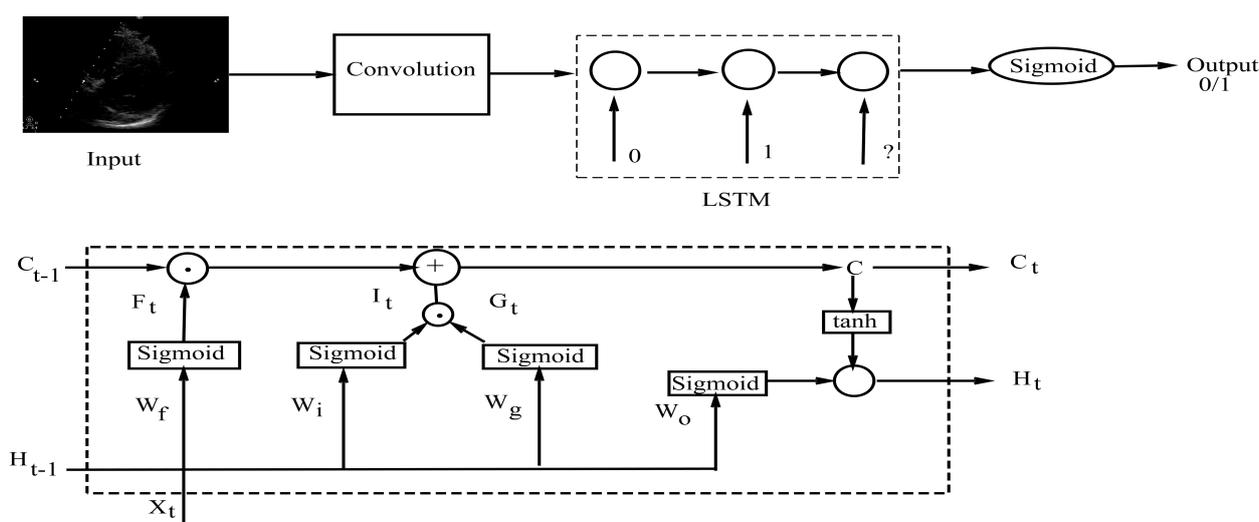


Figure 3. Long Short Term Memory (LSTM) diagram with an overall flowchart of how the model works where an input image is passed to a convolution layer, then it is forwarded to the architecture whereby images are classified based on the previous and present classification.

3.3.2. Variational Autoencoder + Convolutional Neural Network (VAE-CNN)

Combining CNN with other methods help the network excel at spatial relationships [31]. Convolutional layers are a significant building block in deep neural networks. However, the gradient computation of the convolution network remains a challenge in TensorFlow. Many researchers argue that even random convolutions are content [32]. On the other hand, Autoencoder is a powerful generative model that takes CNN idea, which is useful for reconstructing its output through encodings.

The overall diagram of how CNN is combined with VAE is shown in Figure 4. VAE has been taken as a method as it works with a diverse range of data [33]. The overall procedure starts with the input being passed to a convolutional layer with filter size 3×3 and stride of 2×2 followed by another convolutional layer with the same filter size and no stride. Then it is passed to the Maxpooling layer with size 2×2 that helps in reducing the image size which is then passed to a fully connected layer with 4096 nodes. It is then followed by VAE where a first dense layer of 500 is used, followed by another dense layer of 120 and then followed by vector generation of μ and σ , which will produce a sample vector of 30 [33]. μ is mean, and σ is the standard deviation. It is then passed to a classification layer that produces output class using sigmoid and softmax classifiers for two-class and six-class, respectively. It can also be passed to a decoder but was not done, as in our case, our purpose is classification.

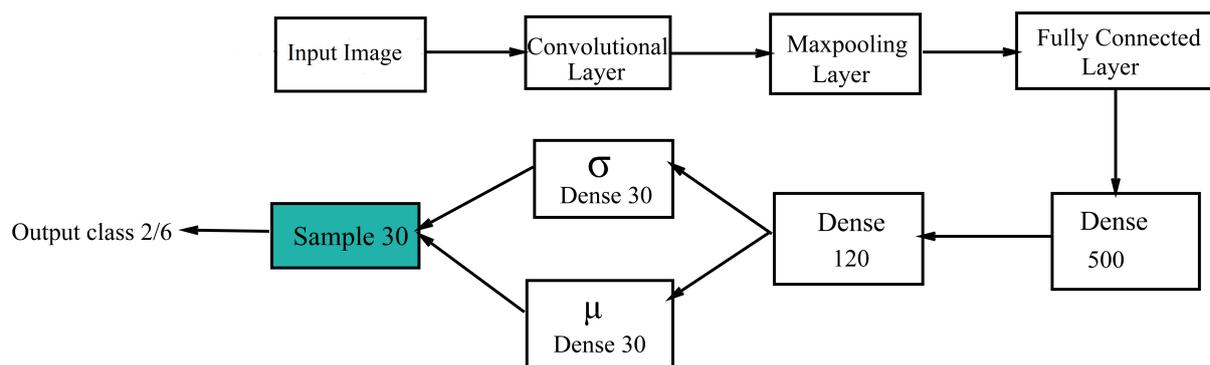


Figure 4. Variational Autoencoder + Convolutional Neural Network (VAE-CNN) flowchart.

3.3.3. Support Vector Machine (SVM) Methodology

SVM is one of the most widely used supervised learning methods for classification in the different fields in medical imaging [14,34–36]. SVM is memory efficient and effective in case of high dimensional spaces. It is not only used for classification but for regression as well. For this paper, the SVM methodology was taken from [14]. Although the paper did not mention the type of SVM being used, but for our purpose, SVM with a linear kernel is used as it is popularly used in many fields. The linear kernel can be represented as [37]:

$$(K(W, I)X, X') = W^T I \quad (7)$$

where $W^T I$ is the sum of the inner product. The output class obtained is 0 or 1, 0 for normal (absence of abnormalities) class, and 1 for abnormal (presence of abnormalities) class.

The difference in parameters used in LSTM and VAE-CNN is provided in Table 1. After a brief discussion on the different methods used in this paper, the next section is provided with experimental results and conclusions.

Table 1. Architecture for LSTM and VAE-CNN.

Method	Layer Number	Layer Name	Layer Properties
LSTM	1	Input layer	Size 224×224
	2	Convolutional layer	3×3 filter size, stride = 2, output size = 115×115
	3	Flatten	36,963
	4	LSTM model	126 units
	5	Dropout	50 % dropout
	6	Fully connected layer	50
	7	Rectified Linear units	Rectified Linear Units
	8	Sigmoid/Softmax	Sigmoid/Softmax
	9	Classification output	2 (normal or abnormal) and 6 (types of regurgitations)
VAE-CNN	1	Input image	224×224
	2	Convolutional layer	3×3 filter size, stride = 2
	3	Rectified Linear Units	Rectified Linear Units
	4	Dropout	50 % dropout rate
	5	Max pooling	stride = 2, output size = 115×115
	6	Dropout	50 % dropout rate
	7	Fully Connected	9075
	8	Fully Connected	500
	9	Fully Connected	100
	10	Fully Connected	Sample vector, 30 (Standard deviation), 30 (Mean)
	11	Fully Connected	30
	12	Sigmoid/Softmax	Sigmoid/Softmax
	13	Classification layer	2 (normal or abnormal) and 6 (types of regurgitations)

4. Experiment and Result Analysis

The implementation was carried out using jupyter notebook, which is a readily available and open-source web application for python programming language and google colab. The result is divided into two parts, firstly, classification into normal or abnormal, and secondly classification into different types of regurgitations. We have used k fold cross validation as our first approach consisting of varieties of folds (2, 5 and 10). The second approach used here is generalization capability in medical diagnostic where there is no observation in the training phase and no data in the training is in the testing phase [38]. If that is not maintained there might be an astonishing relationship between the obtained status and real identity which results in unrealistic results. Here no data in the training phase is in the testing phase as data here are based on patient data where a patient in the training and validation phase (train-test split) is not used in the testing phase (separate CSV file). Except in two patient cases for six-class classifications, the same data was augmented and kept in the same CSV file. The training data is labelled and the testing data is unlabelled. The optimizer used is Adam and batch size of 50.

4.1. Performance Metrics

Performance metrics are used to evaluate and for checking the quality of performance by the algorithms. Accuracy is considered to be one of the most widely applied performance metrics in classification. The different performance metrics used in the paper are as follows:

4.1.1. Classification Accuracy

It is a measure that calculates the ratio of the correct classification to the total number of samples. For class having the same number of samples, accuracy itself is sufficient as a metric.

4.1.2. Logarithmic Loss

It works by castigating false classification. The lower the loss, the better will be the accuracy. It works well for multi-class classification. Here, Binary crossentropy is used as a loss function.

4.1.3. Confusion Matrix

It is a matrix that describes the complete performance of a model. The following can be calculated based on the confusion matrix.

1. Precision

It is the fraction of True Positives (TP) and False Positives (FP). For two class classification, precision can be calculated using

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (8)$$

For six-class classification precision can be calculated using

$$P_i = \frac{\text{TP}_i}{\text{TP}_i + \sum_{j=0}^i E_{ji} + \sum_{j=i+1}^5 E_{ji}} \quad (9)$$

where i ranges from 0 to 5.

2. Recall

It represents the fraction of True Positives (TP) and False Negatives (FN). For two-class classification recall can be calculated as

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (10)$$

For six-class it can be generalized as

$$R_i = \frac{\text{TP}_i}{\text{TP}_i + \sum_{j=0}^{i-1} E_{ij} + \sum_{j=i+1}^5 E_{ij}} \quad (11)$$

3. F1 Score

It is a harmonic mean of precision and recall. For two-class classification it is given by:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

For six-class classification it can be calculated from recall (R_i) and precision (P_i) as:

$$F_i = 2 \times \frac{P_i \times R_i}{P_i + R_i} \quad (13)$$

4.2. Classification into Normal or Abnormal

4.2.1. Dataset

Here, classification into normal or abnormal is carried out using two types of echo images. Obtained data are 2D images in Joint Photographic Experts Group (JPEG) format and 3D color Doppler in BitMap (BMP) format. Data were collected from Hope clinic, Shillong. For the validation phase, 10% of the total data were obtained from the training set used in the training phase. For the testing phase, data are separated in an unused folder (which is later saved in CSV file) where these are tested in the later part after validation. The number of images used for 2D image classification and 3D Doppler image classification is 1070 and 540, respectively. Out of which, 10% is for validation and the rest for training. Excluding these, there are 38 number of 2D images and 10 number of Doppler images for testing purposes. Testing data are separated from training and validation in the experiment for prediction purposes. The total number of 2D images is 1108, and 3D Doppler is 550. For k fold classification the data from all the phases are combined. The k fold cross validation was run multiple times to obtain the same number of data in both the methodologies for plotting the confusion matrix. This is done so that comparison can be made with the same number of data even though the pattern obtained in both cases are different.

4.2.2. Output

Figures 5–14 show the confusion matrix for 2D images and 3D color Doppler and Table 2, the result showing accuracy, precision, recall, and F1 Score. From the graph in Figure 15 plotted for all performance metrics and output obtained without k fold, we can conclude that VAE-CNN gives better output in almost all cases. In the case of SVM, in testing for 2D images it is almost equivalent to that of VAE-CNN. In other cases, deep learning methodologies are better compared to SVM in the classification of heart images. It can also be seen that accuracy is better in VAE-CNN, precision in SVM, recall in VAE-CNN, and F1 Score in VAE-CNN for 2D images in the validation phase. In the case of color Doppler, accuracy is better in VAE-CNN, precision in VAE-CNN, recall in SVM, and F1 Score in VAE-CNN in the validation phase. During prediction (testing phase), accuracy in SVM, precision in LSTM, recall in VAE-CNN, and F1 Score in VAE-CNN and SVM is better in the case of 2D images. Overall, VAE-CNN gives a better output compared to the other two methods. Using k fold cross validation, in case of 2D images VAE-CNN performs better in the case of 2 fold and 5 fold and almost equivalent to the others in case of 10 folds. We can also see that for 3D doppler images VAE-CNN performs better compared to LSTM and SVM in all the three folds. Overall, VAE-CNN gives a better output when using k fold which is in the case of generalization classification as well.

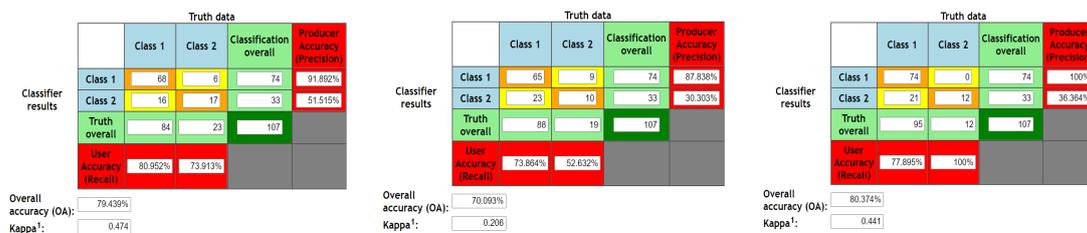


Figure 5. Confusion matrix for validation phase of 2D images for SVM, LSTM and VAE3CNN respectively.

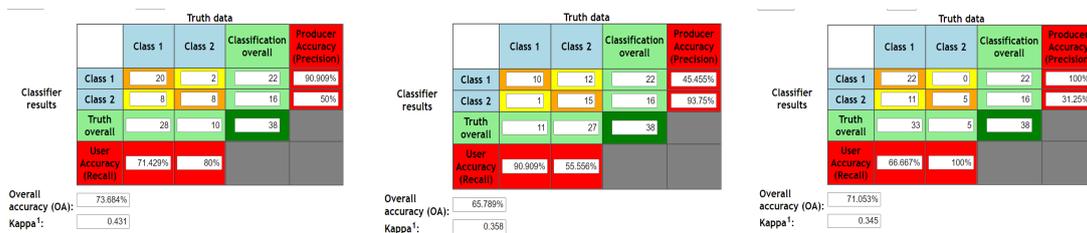


Figure 6. Confusion matrix for testing phase of 2D images for SVM, LSTM and VAE-CNN respectively.

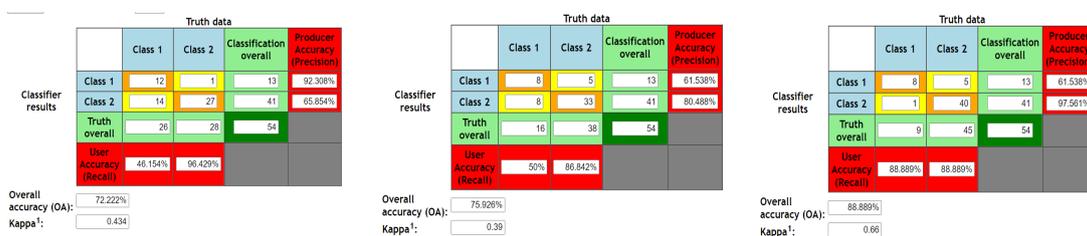


Figure 7. Confusion matrix for Validation phase of 3D Doppler images for SVM, LSTM and VAE-CNN respectively.

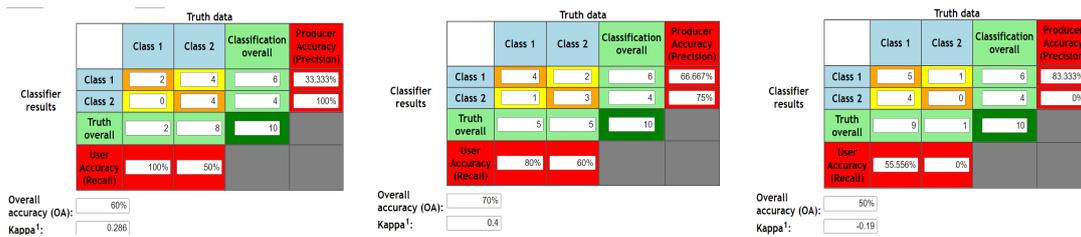


Figure 8. Confusion matrix for testing phase of 3D Doppler images for SVM, LSTM and VAE-CNN respectively.

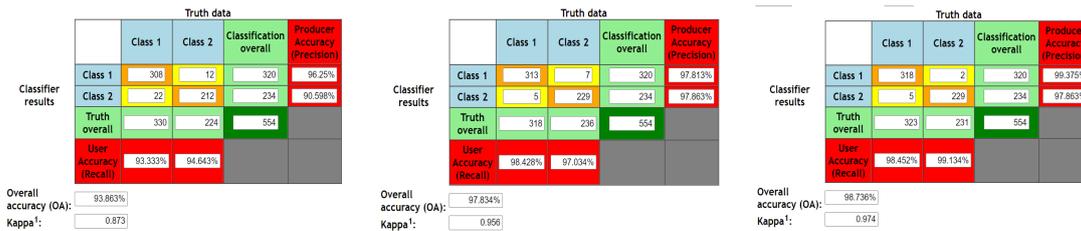


Figure 9. Confusion matrix for 2 fold cross validation of 2D images for SVM, LSTM and VAE-CNN respectively.

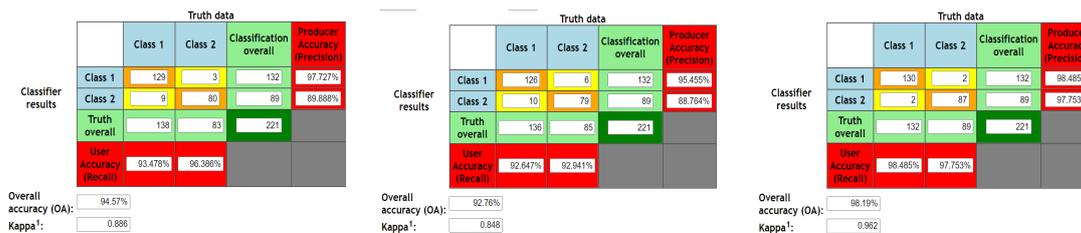


Figure 10. Confusion matrix for 5 fold cross validation of 2D images for SVM, LSTM and VAE-CNN respectively.

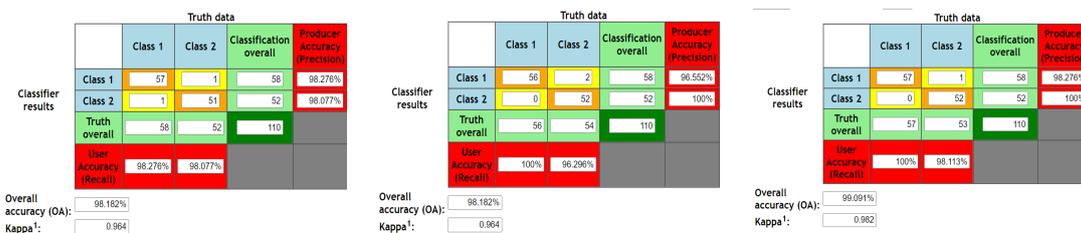


Figure 11. Confusion matrix for 10 fold cross validation of 2D images for SVM, LSTM and VAE-CNN respectively.

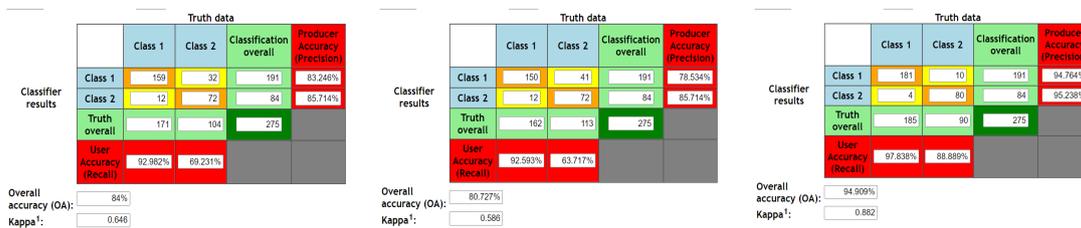


Figure 12. Confusion matrix for 2 fold cross validation of 3D Doppler images for SVM, LSTM and VAE-CNN respectively.

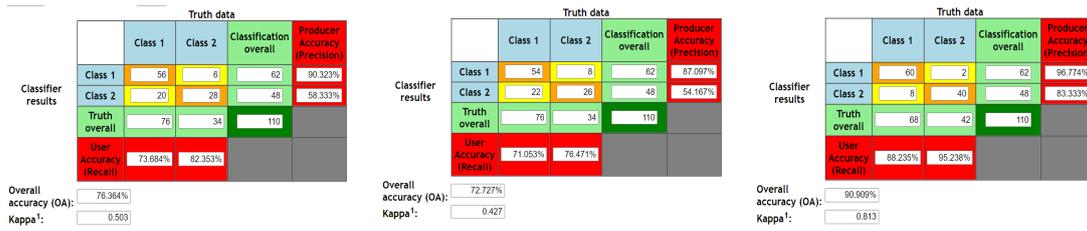


Figure 13. Confusion matrix for 5 fold cross validation of 3D Doppler images for SVM, LSTM and VAE-CNN respectively.

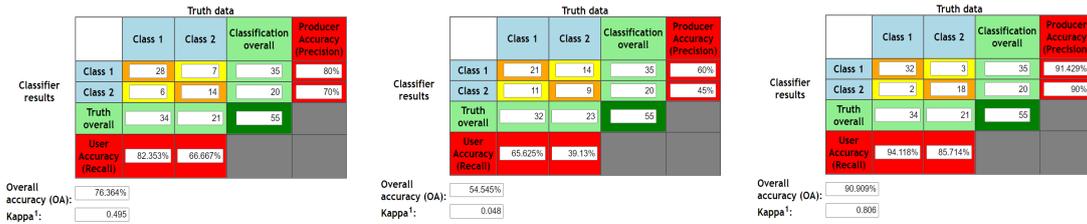


Figure 14. Confusion matrix for 10 fold cross validation of 3D Doppler images for SVM, LSTM and VAE-CNN respectively.

Table 2. Output for 2D images and 3D Doppler images.

Echo Format		3D Doppler Images			2D Images		
Methodologies		LSTM	VAE-CNN	SVM	LSTM	VAE-CNN	SVM
Training Accuracy		0.94	0.96	0.77	0.64	1	0.80
Validation (labelled data)	Accuracy	0.76	0.89	0.72	0.70	0.80	0.79
	Precision	0.50	0.88	0.46	0.66	0.77	0.80
	Recall	0.61	0.61	0.92	0.87	1	0.91
	F1 Score	0.54	0.72	0.60	0.76	0.88	0.85
Testing (unlabelled)	Accuracy	0.70	0.50	0.60	0.47	0.71	0.73
	Precision	0.80	0.55	1	0.90	0.66	0.71
	Recall	0.66	0.83	0.33	0.45	1	0.90
	F1 Score	0.73	0.66	0.49	0.60	0.79	0.79
Testing 2 fold	Accuracy	0.80	0.94	0.84	0.97	0.98	0.94
	Precision	0.92	0.97	0.92	0.97	0.99	0.96
	Recall	0.78	0.94	0.83	0.98	0.98	0.93
	F1 Score	0.84	0.95	0.87	0.97	0.98	0.94
Testing 5 fold	Accuracy	0.72	0.91	0.76	0.92	0.99	0.96
	Precision	0.71	0.88	0.73	0.95	0.98	0.97
	Recall	0.87	0.96	0.90	0.92	0.98	0.94
	F1 Score	0.78	0.91	0.80	0.93	0.98	0.98
Testing 10 fold	Accuracy	0.54	0.92	0.76	0.98	0.98	0.98
	Precision	0.65	0.94	0.80	0.96	0.98	0.98
	Recall	0.60	0.91	0.82	1	1	0.98
	F1 Score	0.62	0.92	0.80	0.97	0.98	0.98

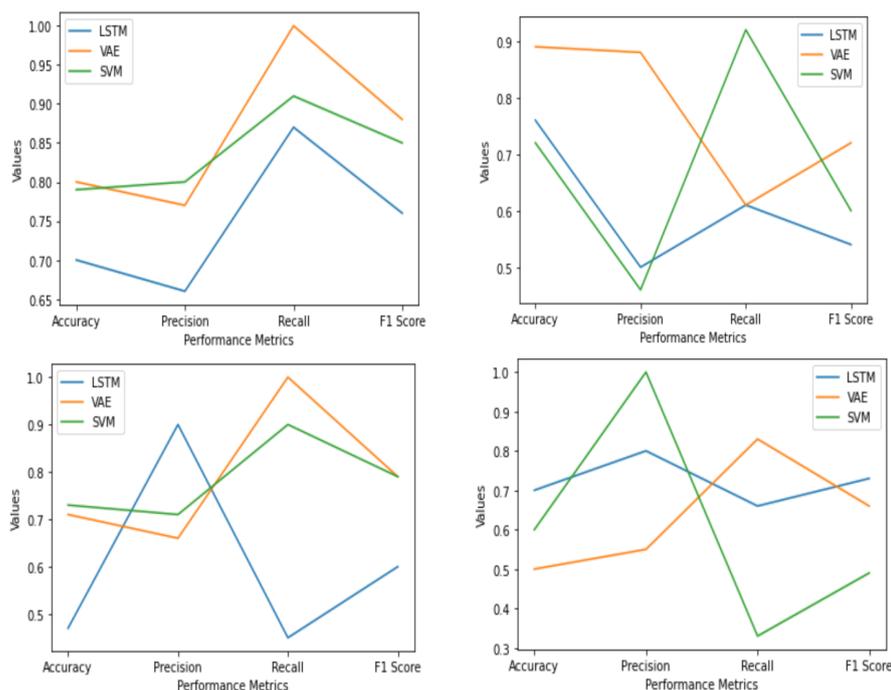


Figure 15. Graph plotted for validation phase of 2D image and color doppler and testing phase of 2D image and color doppler respectively, without *k* fold cross validation.

4.2.3. Statistical Significance Test

Statistical tests are used for comparison of the classifier. Several statistical tests are available out of which paired *T*-test has been used for comparison of LSTM and VAE-CNN to that of existing SVM methodology. As *k* fold cross validation is already considered as a statistical procedure, *T*-test has not been calculated for the same. It is used for determining the mean of two sets, which is equivalent to zero [39]. It shows the significance of a model by specifying the *p*-value obtained from the test. Mathematically it can be calculated using:

$$t = \frac{d}{\sqrt{s^2/n}} \tag{14}$$

where *d* is the mean difference, *s*² is the sample variance and *n* is the number of samples.

Based on the statistical test in Table 3, the result obtained by both deep learning methodologies dominates that of SVM. It is done by considering the value of 0.05. From the total 8 cases, 6 cases show statistically significant improvement compared to SVM. In all cases LSTM and VAE-CNN obtained statistical significance improvement over SVM, though in some cases there is not much significance. Only in 1 case of VAE-CNN and 1 case of LSTM, there was no improvement at all. Using the obtained data, the deep learning methodologies are more effective based on the value of the paired *T*-test. Based on the test of the three methodologies, a difference is observed between the group and therefore is statistically significant.

Table 3. Statistical test for comparing methodologies without *k* fold cross validation.

Type of Image	Methods	Paired <i>T</i> -Test (Validation)	Paired <i>T</i> -Test (Testing)
2D image	SVM vs. LSTM	0.493	0.00014
	SVM vs. VAE-CNN	0.042	0.153
3D Color Doppler	SVM vs. LSTM	0.048	0.0005
	SVM vs. VAE-CNN	0.0003	0.0005

Summary

Some observations based on this classification are as follows:

1. It can be observed that the traditional method could properly classify the different classes as compared to deep learning methodologies.
2. Using different views and types of image format gives an almost equivalent output, which means that these methods work for any view of echo.

4.3. Classification into Types of Regurgitation

Classification for 6 types of regurgitation has been done using videographic images namely, class 0—mitral regurgitation (MR), class 1—aortic regurgitation (AR), class 2—tricuspid regurgitation (TR), class 3—mitral regurgitation, and tricuspid regurgitation (MR+TR), class 4—aortic regurgitation and mitral regurgitation (AR+MR) and class 5—aortic regurgitation, mitral regurgitation and tricuspid regurgitation (AR+MR+TR). These classes were selected based on data availability of the types of regurgitation. Also, classification into two types, i.e., class 0—normal or class 1—abnormal was done using videographic images.

4.3.1. Dataset

The data obtained were in video format, and for each class, 10 patients' data were used. The frame ranges from 33 to 150 for each patient. The total number of images used for two-class classification during training is 2430, where 243 are for validation purposes. For the testing phase, the total number of images is 539. The total number of images used for six-class classification during training is 5160, where 516 are for validation purposes. For the testing phase, the total number of images is 736. In the case of k fold cross validation, the data were combined from all the phases. Here too the methodologies were run multiple times to obtain the same number of data in both the methodologies for plotting the confusion matrix.

4.3.2. Output

Two methods, namely, LSTM and VAE-CNN, are used for comparison purposes. It was not compared with other methods as no such work has been done in the same field using video. These two methods were used to check the applicability of deep learning methodologies in these fields. From the output obtained in Tables 4–6 and graphs in Figures 16 and 17, it could be seen that classification into normal or abnormal gives very accurate results for both the methodologies in the training and validation phase. When it comes to prediction using VAE-CNN, a better result is obtained. Both cases under performed as accuracy during prediction could not be more than 80% in the case of two-class classification. Nonetheless, it can be observed that deep learning methodologies could classify correctly with the highest of 100% accuracy during validation and 95% accuracy during testing/prediction in six-class classification using LSTM. It can also be observed from the output in Figure 18, Tables 6–8, and graphs plotted in Figures 19 and 17 that classification into 6 types of regurgitation give better accuracy and overall performance than that of normal or abnormal classification. This is due to the higher similarities pattern between classes. Few examples showing classification into six-class regurgitation can be seen in Figure 20.

Using k fold cross validation it can be observed that LSTM performed better compared to VAE-CNN in both cases (normal or abnormal and types of regurgitation classification). In the case of two-class classification LSTM gives a 100% accuracy score in all three folds while VAE-CNN gives 99% and 98% unlike in the generalization approach. In the case of six-class classification VAE-CNN gives a different result than the generalization approach, where the accuracy is improved and with the best accuracy of 86%. VAE-CNN did not under perform, but could not overtake LSTM in either of the cases.

Table 4. Confusion matrix for two-class (normal or abnormal) classification.

Methodologies		LSTM		VAE-CNN	
		0	1	0	1
Validation phase	0	132	0	132	0
	1	0	111	0	111
Testing phase	0	191	52	243	0
	1	220	76	191	105
Testing 2 fold	0	812	0	805	7
	1	0	672	5	667
Testing 5 fold	0	323	0	321	2
	1	0	270	1	269
Testing 10 fold	0	182	0	181	1
	1	0	114	2	112

Table 5. Confusion matrix for validation and testing into six-class (type of regurgitation) classification.

Methodologies		LSTM						VAE-CNN					
		0	1	2	3	4	5	0	1	2	3	4	5
Validation Phase	0	73	3	0	5	0	0	77	0	0	4	0	0
	1	5	83	3	8	0	0	0	90	1	8	0	0
	2	0	4	80	9	0	0	0	11	82	0	0	0
	3	4	2	3	74	0	0	0	0	6	77	0	0
	4	3	0	4	10	62	0	6	2	4	3	64	0
	5	13	6	6	0	0	56	0	8	6	8	0	59
Testing Phase	0	89	4	0	3	0	0	20	9	67	0	0	0
	1	3	178	2	1	5	0	45	80	22	30	10	2
	2	0	0	134	0	0	0	15	10	95	14	0	0
	3	0	0	0	126	0	0	0	48	0	56	22	0
	4	0	0	2	12	87	0	0	60	0	12	29	0
	5	0	0	0	0	0	91	0	23	36	15	0	17

Table 6. Output for Video 2 (normal or abnormal) and six-class (types of regurgitation).

Echo Format		Video 2 Class		Video 6 Class	
		LSTM	VAE-CNN	LSTM	VAE-CNN
Training Accuracy		1	1	0.86	0.93
Validation	Accuracy	1	1	0.85	0.90
	Precision	1	1	0.84	0.88
	Recall	1	1	0.82	0.86
	F1 Score	1	1	0.79	0.87
Testing	Accuracy	0.49	0.64	0.95	0.39
	Precision	0.46	0.55	0.95	0.44
	Recall	0.78	1	0.95	0.37
	F1 Score	0.56	0.70	0.94	0.37
Testing 2 fold	Accuracy	1	0.99	0.88	0.85
	Precision	1	0.99	0.89	0.85
	Recall	1	0.99	0.89	0.88
	F1 Score	1	0.99	0.89	0.86
Testing 5 fold	Accuracy	1	0.99	0.87	0.85
	Precision	1	0.99	0.86	0.86
	Recall	1	0.99	0.89	0.87
	F1 Score	1	0.99	0.87	0.86
Testing 10 fold	Accuracy	1	0.98	0.89	0.86
	Precision	1	0.98	0.89	0.86
	Recall	1	0.98	0.89	0.89
	F1 Score	1	0.98	0.89	0.87

Table 7. LSTM precision, recall and F1 Score for validation and testing phase for six-class classification.

No	Precision	Validation	Testing	Recall	Validation	Testing	F1 Score	Validation	Testing
1	P0	0.74	0.96	R0	0.90	0.92	F0	0.70	0.93
2	P1	0.82	0.97	R1	0.83	0.94	F1	0.82	0.95
3	P2	0.83	0.97	R2	0.86	1	F2	0.84	0.98
4	P3	0.69	0.88	R3	0.89	1	F3	0.77	0.93
5	P4	1	0.95	R4	0.78	0.86	F4	0.87	0.90
6	P5	1	1	R5	0.66	1	F5	0.79	1

Table 8. VAE-CNN precision, recall and F1 Score for validation and testing phase for six-class classification.

No	Precision	Validation	Testing	Recall	Validation	Testing	F1 Score	Validation	Testing
1	P0	0.92	0.25	R0	0.95	0.20	F0	0.94	0.22
2	P1	0.81	0.34	R1	0.90	0.42	F1	0.85	0.36
3	P2	0.82	0.43	R2	0.88	0.70	F2	0.84	0.53
4	P3	0.77	0.44	R3	0.92	0.44	F3	0.89	0.44
5	P4	1	0.47	R4	0.81	0.28	F4	0.89	0.37
6	P5	1	0.89	R5	0.72	0.18	F5	0.84	0.30

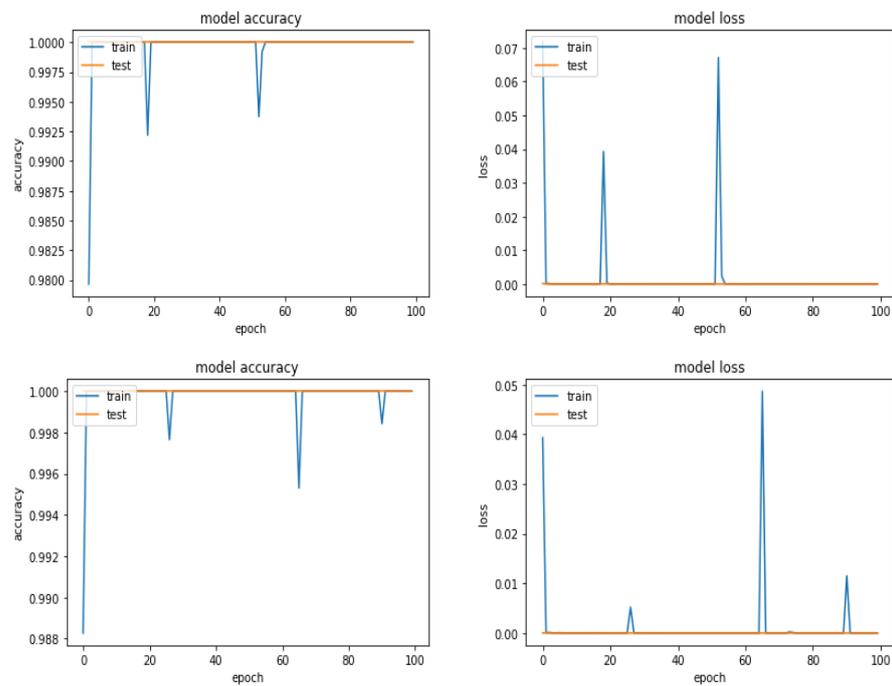


Figure 16. Accuracy and Loss for LSTM and VAE-CNN for two-class classification for training and validation phase respectively.

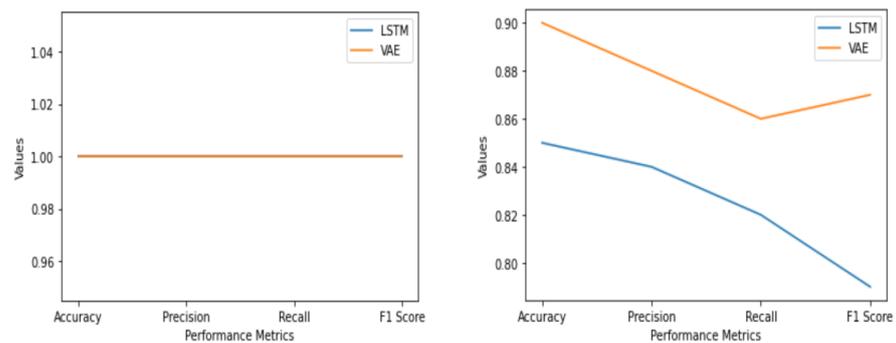


Figure 17. Cont.

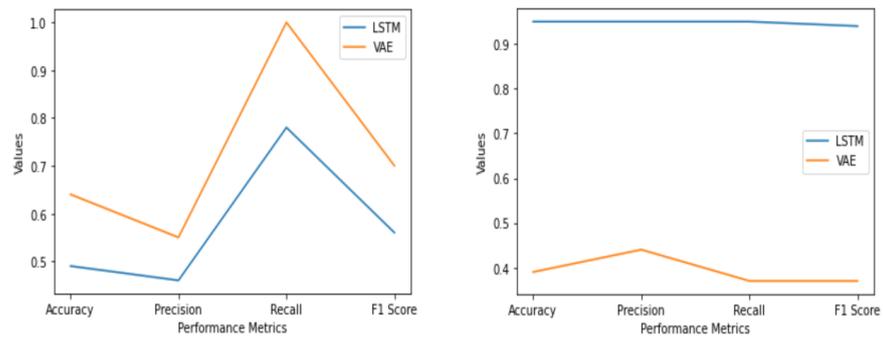


Figure 17. Validation phase for two-class and six-class and testing phase for two-class and six-class respectively.

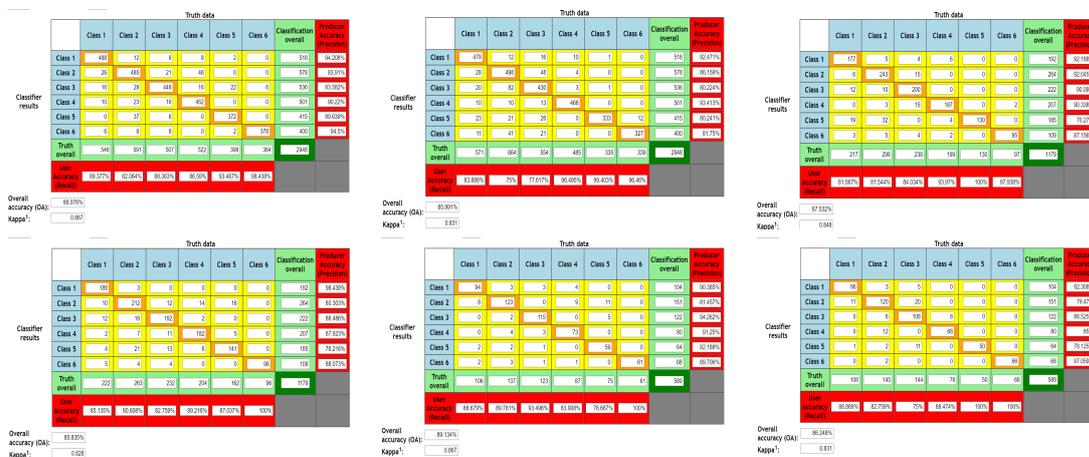


Figure 18. Confusion matrix for 2-fold, 5-fold and 10-fold cross validation for LSTM and VAE-CNN, respectively.

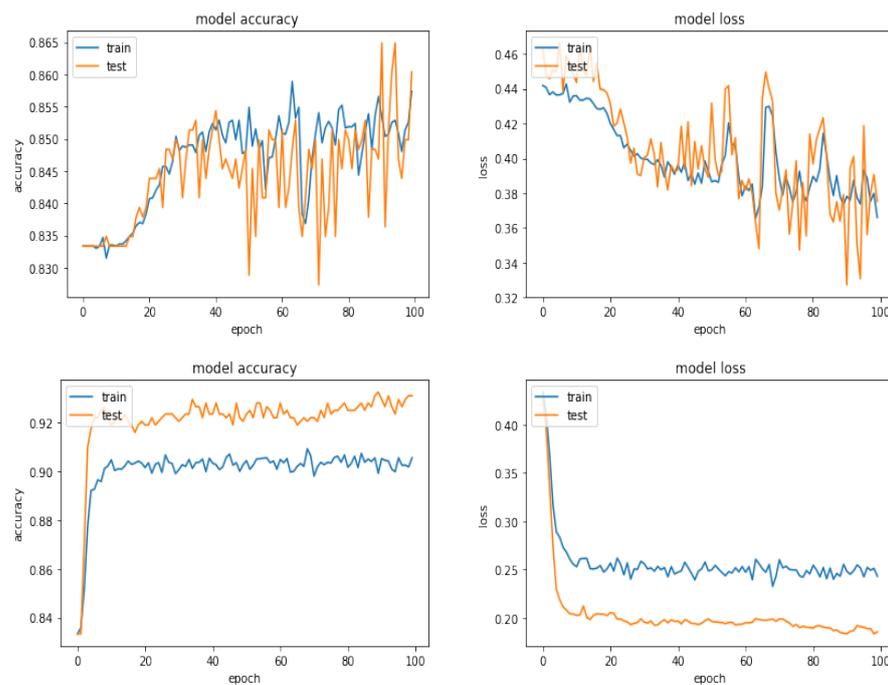


Figure 19. Accuracy and Loss for LSTM and VAE-CNN for six-class classification for training and validation phase, respectively.



Figure 20. Output of expected and predicted class for six-class classification using generalization approach.

Summary

It could be observed that accuracy, precision, recall, and F1 score is better in the validation phase for two-class classification than that of six-class classification. Another output could be observed in the testing phase (generalization approach) where LSTM is found to be better than VAE-CNN using color Doppler for six-class classification. In the testing phase (generalization approach), the output of VAE-CNN is better than of LSTM for the two-class classification. Low accuracy can occur due to fewer data used. Overall, deep learning can be applied and used instead of the SVM method.

With the number of images increases, VAE performance too increases. For all other cases, the pattern of output is similar or the same except in six-class classification where LSTM performs better in case of testing using generalization approach (train test split and prediction). This is due to inadequate patient data available where few repetitive data had to be used for two different classes. This causes misclassification where a patient having both AR and MR is treated to have only MR or AR. This shows that VAE cannot classify properly for classes having two abnormalities in the same frame. LSTM performs better using video images compared to VAE-CNN. LSTM performs better since it is an RNN based method, it has time as a factor that can predict what is the next class based on the present and previous inputs. LSTM sometimes fails to classify images of the different class taking the previously obtained classes as the next class. Varying encoding on every single pass makes VAE-CNN difficult to classify frames of the same class which is its disadvantage. However, VAE-CNN has both the property of CNN and VAE which provides a continuous latent space and made interpolation simpler and sampling easier. In conclusion, we can say VAE-CNN gives better output using static images and LSTM using videographic images from k fold cross validation and generalization approach.

5. Conclusions and Future Work

Heart abnormalities classification has been little explored in the field of cardiology. It is an important aspect for detecting any future diseases. Any step that makes diagnosis more accessible and a tool in the future for human intervention can never be considered vague. Several works have been done in the past, but using deep learning methodologies

or machine learning models has not been explored much in this field. This paper presented two such methodologies in quest of a better algorithm that can better classify the types of regurgitation and class consisting of abnormalities and without any abnormalities. From the obtained output, it can be concluded that using deep learning methodologies regurgitation can be better classified as compared to a well-known SVM method. Using LSTM and VAE prove to be an efficient and effective way in abnormalities detection where the accuracy in most cases is high. Using such algorithms provide a solution to a cardiologist and ease the process of diagnosis. This will reduce human effort and can be used in early detection and for better diagnosis. In this paper, we have used clinical data, not process data, which could be the reason for lesser than expected accuracy in some cases. Work can be done to achieve a greater number of properly processed data in the future using the videographic echo with keyframe extraction and segmentation. This paper is an initiative for the application of deep learning in such kind of works, which can be further expanded. More experiments are needed for a better diagnosis that can ease and even replace human exertion to some extent.

Author Contributions: Conceptualization, I.W. and A.K.M.; methodology, I.W.; software, I.W.; validation, A.K.M. and G.S.; formal analysis, I.W.; investigation, I.W.; writing—original draft preparation, I.W.; writing—review and editing, A.K.M., M.J. and P.C.; supervision, A.K.M., G.S., P.C., M.J., Z.L. and E.J.; Funding acquisition, Z.L. and E.J.; project administration, A.K.M. and G.S. All authors have read and agreed to the published version of the manuscript.

Funding: Publication of this article was financially supported by the Chair of Electrical Engineering, Wrocław University of Science and Technology.

Data Availability Statement: Limited data available on request due to the large size of the data.

Acknowledgments: A special appreciation and thanks go to D. S. Sethi, Director of Hope clinic, Shillong, India for providing the data and for evaluation and identification of the different types of abnormalities.

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

References

1. MayoClinic: Heart Disease. Available online: <https://www.mayoclinic.org/diseases-conditions/heart-disease/diagnosis-treatment/drc-20353124> (accessed on 7 April 2019).
2. Phoenixheartcenter: Echocardiograms: Transthoracic (TTE) & Transesophageal (TEE). Available online: <http://www.phoenixheartcenter.com/echocardiograms-tte-tee/> (accessed on 7 April 2019).
3. Wikipedia: Echocardiography. Available online: <https://en.wikipedia.org/wiki/Echocardiography> (accessed on 5 April 2019).
4. Echocardiography for Emergency Physicians. Available online: <https://www.acep.org/sonoguide/cardiac.html> (accessed on 5 April 2019).
5. Transthoracic and Transesophageal Echocardiography (TEE) and Stress-Echo. Available online: <https://www.tcavi.com/services/transthoracic-and-transesophageal-echocardiography-and-stress-echo/> (accessed on 5 March 2020).
6. Rowin, E.J.; Maron, B.J.; Haas, T.S.; Garberich, R.F.; Wang, W.; Link, M.S.; Maron, M.S. Hypertrophic cardiomyopathy with left ventricular apical aneurysm: Implications for risk stratification and management. *Am. Coll. Cardiol.* **2017**, *69*, 761–773. [CrossRef] [PubMed]
7. Normal Echocardiographic View & Anatomy. Available online: <http://www.ksecho.org=workshop=2017fallw=file=program=26%20Normal.pdf> (accessed on 5 April 2019).
8. Sonography Resources. Available online: <https://sites.austincc.edu/sonography-resources/> (accessed on 13 February 2021).
9. Lancellotti, P.; Tribouilloy, C.; Hagendorff, A.; Popescu, B.A.; Edvardsen, T.; Pierard, L.A.; Badano, L.; Zamorano, J.L. Recommendations for the echocardiographic assessment of native valvular regurgitation: An executive summary from the European association of cardiovascular imaging. *Eur. Heart J. Cardiovasc. Imaging* **2013**, *14*, 611–644. [CrossRef] [PubMed]
10. Perloff, J.K.; Marelli, A. *The Clinical Recognition of Congenital Heart Disease*, 3rd ed.; Jaypee Brothers: New Delhi, India, 2015; pp. 166–678.
11. Gera, R.M. *Step by Step Pediatric Echocardiography*, 3rd ed.; Jaypee Brothers: New Delhi, India, 2015; pp. 10–80.
12. Healthy Living. Cardiac Catheterization. Available online: <https://www.heart.org/en/health-topics/heart-attack/diagnosing-a-heart-attack/cardiac-catheterization> (accessed on 9 March 2020).
13. Family Practice Notebook: Parasternal Long-Axis Echocardiogram View. Available online: <https://fpnotebook.com/cv/Rad/PrstrnlLNgAxsEchcrdgrmVw.htm> (accessed on 5 March 2020).

14. Allan, G.; Nouranian, S.; Tsang, T.; Seitel, A.; Mirian, M.; Jue, J.; Hawley, D.; Fleming, S.; Gin, K.; Swift, J.; et al. Simultaneous analysis of 2d echo views for left atrial segmentation and disease detection. *IEEE Trans. Med. Imaging* **2017**, *36*, 40–50. [[CrossRef](#)] [[PubMed](#)]
15. Afshin, M.; Ayed, I.B.; Punithakumar, K.; Law, M.; Islam, A.; Goela, A.; Peters, T.; Li, S. Regional assessment of cardiac left ventricular myocardial function via MRI statistical features. *IEEE Trans. Med. Imaging* **2013**, *33*, 481–494. [[CrossRef](#)] [[PubMed](#)]
16. Diabetes/Kidney/Heart Disease. *Cardiology Clinics*. Available online: [https://www.cardiology.theclinics.com/article/S0733-8651\(19\)30037-2/pdf](https://www.cardiology.theclinics.com/article/S0733-8651(19)30037-2/pdf) (accessed on 5 April 2019).
17. Balaji, G.N.; Subashini, T.S.; Chidambaram, N. Automatic classification of cardiac views in echocardiogram using histogram and statistical features. *Procedia Comput. Sci.* **2015**, *46*, 1569–1576. [[CrossRef](#)]
18. Balaji, G.N.; Subashini, T.S.; Suresh, A. An efficient view classification of echocardiogram using morphological operations. *J. Theor. Appl. Inf. Technol.* **2014**, *67*, 732–735.
19. Nandagopalan, S. Efficient and Automated Echocardiographic Image Analysis through Data Mining Techniques. Ph.D. Thesis, Amrita Vishwa Vidyapeetham University, Amritanagar, India, 2012.
20. Pinjari, A.K. Image Processing Techniques in Regurgitation Analysis. Available online: <http://shodhganga.inflibnet.ac.in:8080/jspui/handle/10603/10102> (accessed on 25 July 2013).
21. Baumgartner, H.; Hung, J.; Bermejo, J.; Chambers, J.B.; Edvardsen, T.; Goldstein, S.; Lancellotti, P.; LeFevre, M.; Miller, F., Jr.; Otto, C.M. Recommendations on the echocardiographic assessment of aortic valve stenosis: A focused update from the European association of cardiovascular imaging and the American society of echocardiography. *Eur. Heart J. Cardiovasc. Imaging* **2016**, *2318*, 254–275.
22. Strunic, S.L.; Rios-Gutiérrez, F.; Alba-Flores, R.; Nordehn, G.; Bums, S. Detection and classification of cardiac murmurs using segmentation techniques and artificial neural networks. In Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining, Honolulu, HI, USA, 1–5 April 2007; pp. 397–404.
23. Samaneh, M.; Wirza, R.; Sulaiman, P.S.; Dimon, M.Z.; Khalid, F.; Tayebi, R.M. Segmentation methods of echocardiography images for left ventricle boundary detection. *J. Comput. Sci.* **2015**, *11*, 957–970.
24. Naing, O.Y.; Khaing, A.S. Left ventricle segmentation from heart echo images using image processing techniques. *Int. J. Sci. Eng. Technol. Res.* **2014**, *3*, 1606–1612.
25. Aljanabi, M.; Qutqut, H.M.; Hijjawi, W. Machine learning classification techniques for heart disease prediction: A review. *Int. J. Eng. Technol.* **2018**, *7*, 5373–5379.
26. Alarsan, F.I.; Younes, M. Analysis and classification of heart diseases using heartbeat features and machine learning algorithms. *J. Big Data* **2019**, *6*, 1–15. [[CrossRef](#)]
27. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1738. [[CrossRef](#)] [[PubMed](#)]
28. Long Short-Term Memory. Available online: <http://axon.cs.byu.edu/martinez/classes/778/Papers/lstm.pdf> (accessed on 10 April 2019).
29. Recurrent Neural Networks. Available online: http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture10.pdf (accessed on 10 April 2019).
30. Long Short-Term Memory. Available online: <https://towardsdatascience.com/understanding-lstm-and-its-quick-implementation-in-keras-for-sentiment-analysis-af410fd85b47> (accessed on 10 April 2019).
31. Sequence Classification with LSTM Recurrent Neural Networks in Python with Keras. Available online: <https://machinelearningmastery.com/sequence-classification-lstm-recurrent-neural-networks-python-keras/> (accessed on 10 April 2019).
32. Martin, A.; Chu, A.; Goodfellow, I.; McMahan, H.B.; Mironov, I.; Talwar, K.; Zhang, L. Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016; pp. 308–318.
33. Intuitively Understanding Variational Autoencoders. Available online: <https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf> (accessed on 9 April 2019).
34. Taie, S.; Ghonaim, W. CSO-based algorithm with support vector machine for brain tumor’s disease diagnosis. In Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Kona, HI, USA, 13–17 March 2017; pp. 183–188.
35. Balasubramanian, C.; Sudha, B. Comparative study of de-noising, segmentation, feature extraction, classification techniques for medical images. *Int. J. Innov. Res. Sci. Eng. Technol.* **2014**, *3*, 1194–1199.
36. Nelly, G.; Montseny, E.; Sobrevilla, P. State of the art survey on MRI brain tumor segmentation. *Magn. Resonance Imaging* **2013**, 1426–1438.
37. Intelligent Systems: Reasoning and Recognition. Available online: <https://ensimag.grenoble-inp.fr/fr/formation/intelligent-systems-reasoning-and-recognition-4mmsirr6> (accessed on 11 February 2019).
38. Badža, M. M.; Barjaktarović, M.C. Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. *Appl. Sci.* **2020**, *10*, 1999. [[CrossRef](#)]
39. T-test using Python and Numpy. Available online: <https://towardsdatascience.com/inferential-statistics-series-T-test-using-numpy-2718f8f9bf2f> (accessed on 10 February 2019).