International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 6 Issue 1, January - 2022, Pages:7-23

Heart Sounds Analysis and Classification for Cardiovascular Diseases Diagnosis using Deep Learning

Mohammad Khaleel Alnajjar and Samy S. Abu-Naser

Department of Information Technology,
Faculty of Engineering & Information Technology,
Al-Azhar University-Gaza, Palestine

***Abstract**: Heart sounds are created during the cardiac cycle, blood flows into the heart chambers as the cardiac valves open and close. The blood flow creates auditory sounds; the more turbulent the blood flow, the more vibrations are produced. In healthy adults, there are two normal heart sounds that occur in sequence with each heartbeat. These are the first heart sound (S1) and second heart sound (S2), produced by the closing of the atrioventricular valves and semilunar valves, respectively. The cardiovascular diseases (CVDs) are the number one cause of death globally: more people die annually from CVDs than from any other cause. An estimated 7.2 million deaths were due to coronary heart diseases. This study is developed to produce a deep learning model to detect signs of heart diseases through classifying heart sounds. The anticipated method would serve as an initial screening of cardiac diseases which can help in detecting signs of heart diseases. The results of the screening can be used by pathologies of both a hospital environment by a doctor (using a digital stethoscope) and at home by the patient (using a mobile device). Our classification Heart Sounds model uses Deep Learning (DL) techniques using spectrogram by converting audio data to images that utilize the advantages of Mel-Frequency Cepstrums (MFC) to extract perceptual features Mel Frequency Cepstral Coefficient (MFCC). The results improved results and a significant reduction amount of training and testing times using evaluation metric including Accuracy, Loss, Precision, Recall, and F1-Score against including VGG16, ResNet, MobileNet, Inception V3 and Xception. The proposed model attained 100% F1-score accuracy and 100% testing accuracy.*

Keywords: Hear sounds, cardiac diseases, diagnosis, deep learning, and classification.

1. Introduction

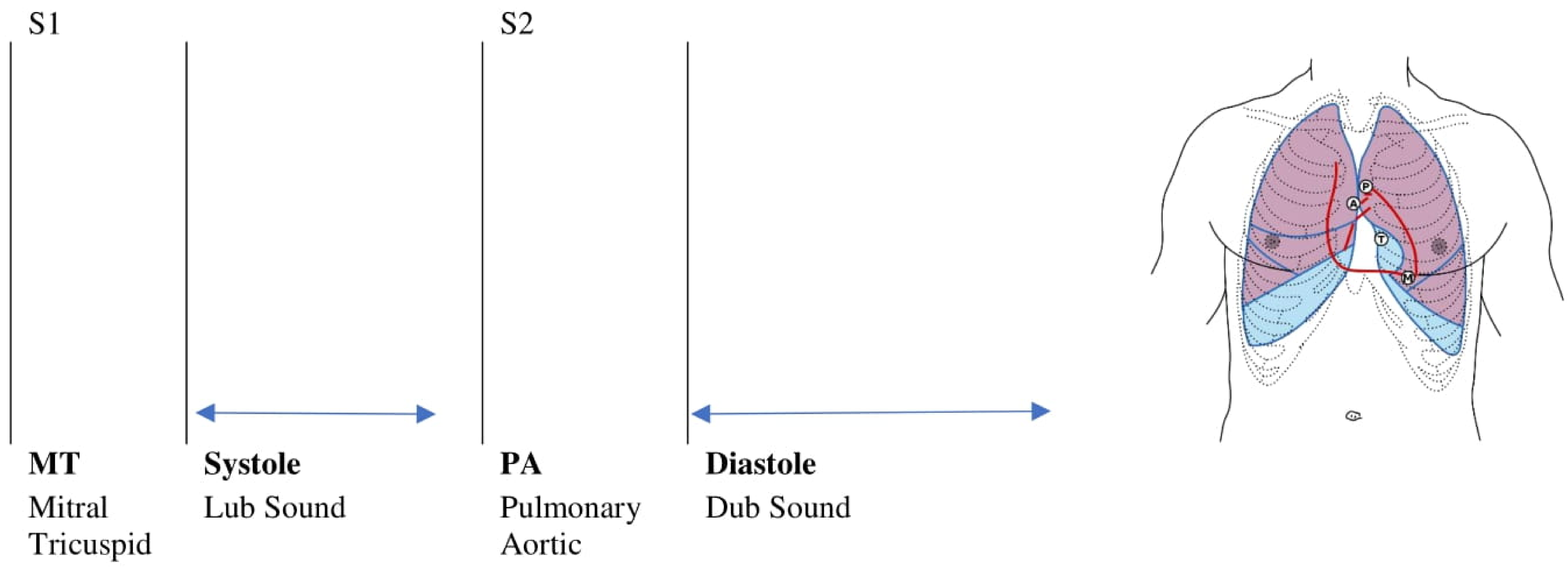
Computers can replicate and modify human-like behavior through machine learning. Each interaction, each action completed, becomes something that the system can learn and utilize as knowledge for the next time using machine learning. [1] Study computer models related to learning processes contribute the subject machine learning multiple in so many fields. [2] Machine learning has always been an integral part of artificial intelligence, and its methodology has grown learning approaches and knowledge representations and applications. ML utilizes the availability of big data and hardware in a clusters, and machines across multiple computational devices, including multicore CPUs, general-purpose GPUs, and custom designed ASICs known as Tensor Processing Units (TPUs)[3].

In recent years, deep artificial neural networks have competed and won many competitions in pattern recognition and machine learning [4]. In order to learn complicated functions and algorithms in higher abstractions such as computer vision, language processing, and related AI disciplines, deep architectures consists of multiple levels of non-linear operations are necessity [5]. In supervised learning, classification algorithms is capable of categorizing data based on past knowledge to predict the labels of future instances [4, 6]. On the other hand, unsupervised learning of representations has been found useful where there are many unlabeled examples and few labeled ones (semi-supervised learning) in which algorithms have successfully been used to learn a hierarchy of features, i.e., to yield representations. and discover structures in large data sets [7] [8].

Using Mel spectra and raw waveform utilizes convolutional neural networks in audio recognition is a central feature in deep learning models [9] [10] [11].

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 6 Issue 1, January - 2022, Pages:7-23

1.1 Hear Sounds and its relation to Cardio Vascular Diseases



Research suggests heart sounds are directly linked to cardiovascular health and diseases which are mainly caused by insufficient blood circulation [12]. The study of the heart functions is significant for the diagnosis of several cardiac causes and effects using the four cardiac sound noises (S1, S2, S3 and S4) which are correlated with cardiac activity and gives information about the state of the heart [13] [14]. In general, heart sound sounds and signals, which are formed by vibrations caused by cardiovascular activities such as heart contractions, heart valve closure, and compression of the ventricular wall. Those signals are transmitted into the tissues to the surface of the human chest and can be perceived by the auditory systems and recorded with electronic devices [15] [16]. Using the electronic stethoscope which provides the digital recording of the heart sound called phonocardiogram (PCG) to extract heart sounds and the methodology for diagnosis of cardiac dysfunctions based on deep learning classification techniques will help trained and untrained medical professionals to facilitate its use in medical fields [17] [18] [19].

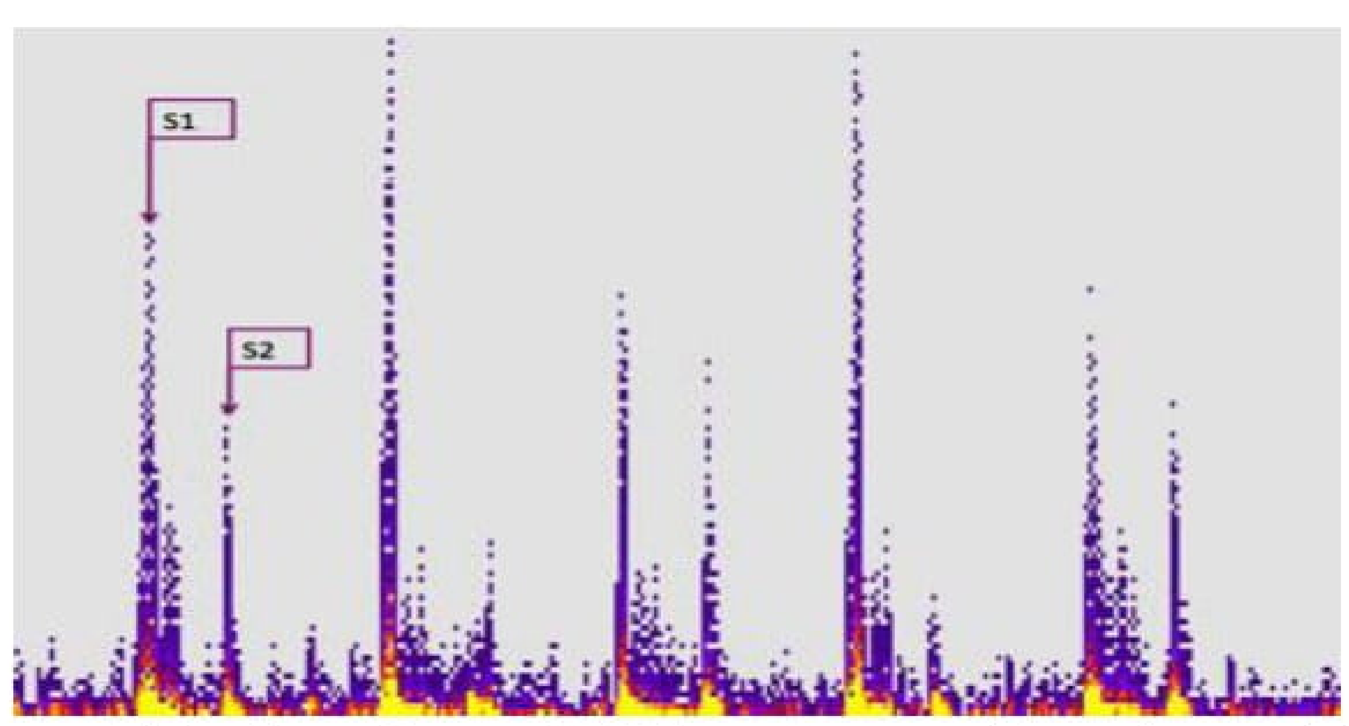


Figure 1 First heart sound (S1) and second heart sound (S2)

1.2 Problem Statement

Nearly 18.6 million people died of cardiovascular disease worldwide in 2019, the most recent year for which figures are available. This represents a 17.1 percent rise over the previous ten years. In 2019, there have been more than 523.2 million cases of cardiovascular disease, a 26.6 percent rise from 2010 [20]. The problem is of particular interest to deep learning involves classification of audio data, usually described as a lub and a dub sounds, to distinguish between classes of interest.
Data can be gathered in real-world situations and frequently which usually contain background noise of every conceivable type caused by the surrounding environment and detector operation [21]. Differences in heart sounds that lead to different heart symptoms can be very small and difficult to distinguish [22]. Successfully classifying this form of audio data, as challenging problem, requires the use of incredibly sophisticated classifiers [23]. Despite the medical importance of cardiovascular problems, deep learning for detection of heart sounds is still a novel research application [24].

1.3 Objectives of the study

The objective of this study is to develop and use deep learning classification models based on Spectrogram Generation. Classification of audio records is developed using state-of-the-art Deep Learning methodologies. The classification of normal vs. various nonnormal heart sounds is proposed to classify heart audio into one of the following categories: Normal, Murmur, Extrasystole, Artifact, Extra Heart Sound.

2.1 Classification of Heart Sounds and Cardio Vascular Diseases

Several researches have been performed in recent years to use audio signal processing to automatically differentiate normal heart sounds from heart beats with abnormal sounds[12]. As for cardiac diseases, a concept of deep learning methodology to automatically classify heart sounds system would be the solution to CVDs diagnosis. [25] Proper distinction of the irregularities in heart sounds can produce results that can to diagnosis cardiac conditions [26]. With appropriate algorithms with deep learning techniques distinguishing heart diseases can be developed. In our research a classification model for heart sounds, which would be used in diagnostic of CVDs, is presented. Nevertheless, imperative advances in the development of different algorithms for classification and analysis of heart sounds, the validity of different approaches has been investigated in our research.

2.2 Audio

The audio files are of varying lengths, between 1 second and 30 seconds. Most information in heart sounds is contained in the low frequency components, with noise in the higher frequencies which is usually suppressed by applying a low-pass filter at 195 Hz. Fast Fourier Transforms (FFT) can provide valuable data about volume and frequency over time.
Time and frequency features are proposed and extracted from heart sounds signals to build and use existing and our proposed deep learning model.

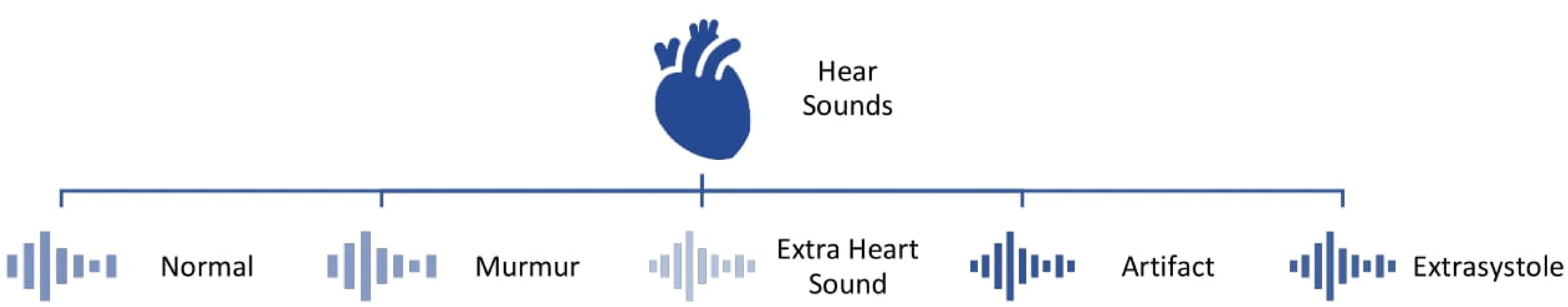


Figure 2 Categories of Heart Sounds

The fundamental heart sounds (FHSs) include the first (S1), second (S2),

2.2.1 Normal Category

In the Normal category represents healthy heart sounds. The normal sound has a clean “lub dub, lub dub” pattern. The “lub” to “dub” timie is shorter than the time coming from “dub” to the subsequent “lub”:

…lub……….dub……………. lub……….dub……………. lub……….dub……………. lub……….dub…

In medicine field we name the lub sound "S1" and the dub sound "S2". Utmost normal heart rates at relaxation will be among about 60 and 100 beats (‘lub dub’s) per minute.

2.2.2 Murmur Category

Heart murmurs can be an indication of numerous heart disorders. The sound comprises a “whooshing, rumbling, roaring, or turbulent fluid” and a “lub” and a “dub” sounds escorted with noise in one of two temporal places:
(1) Among “lub” and “dub”, or
(2) Among “dub” and “lub”.
The misperception with non-medical people that murmurs happens among lub and dub or among dub and lub; not on lub and not on dub. Underneath, you can see an asterisk* at the sites a murmur may be.

…lub..****...dub……………. lub..****..dub ……………. lub..****..dub ……………. lub..****..dub …

or

…lub……….dub…******….lub………. dub…******….lub ………. dub…******….lub ……….dub…

2.2.3 Extra Heart Sound Category

More heart sounds are important signs of disease that can be recognized though the presence of an extra sound for example "lub-lub dub" or a "lub dub-dub". The importance of identify this class is that it cannot be identified by other methods such as ultrasound. Underneath, notice the temporal explanation of the additional heart sounds:

…lub.lub……….dub……….……. lub. lub……….dub………….lub.lub..…….dub…….

or

…lub………. dub.dub…………….lub.……….dub.dub…………….lub……..…….dub. dub……

2.2.4 Artifact Category

Artifact class is utmost dissimilar from the others which covers a wide range of diverse sounds, counting feedback echoes and squeals, music, noise, and speech. There are typically no apparent heart sounds, and therefore slight or no temporal periodicity at frequencies under 195Hz.

2.2.5 Extrasystole Category

Extrasystole sounds might look infrequently and can be recognized since there is a heart sound that is beyond rhythm including extra or skipped heartbeats, e.g. a "lub-lub dub" or a "lub dub-dub". (This is not similar as an additional heart sound as the event is not frequently happening.) An extrasystole might not be a sign of illness. It can occur usually in a mature people and can be common among kids. Yet, in some circumstances extrasystoles can be produced by heart illnesses. If these illnesses are identified earlier, then action is expected to be significantly effective. The temporal explanation of the additional heart sounds:

…........lub……….dub……..………. lub. ………..…….dub…………….lub.lub……..…….dub…….

or

…lub………. dub......…………………lub.…………….dub.dub………………….lub……..…….dub.……

2.2 Mel Scale

The Mel (comes from the word melody) Scale is a logarithmic using natural logarithm (ln of base 10) transformation of a signal's frequency from the Hertz scale to the Mel-Scale is the next

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 6 Issue 1, January - 2022, Pages:7-23

$$m = 1127 \cdot \log\left(1 + \frac{f}{700}\right)$$

Equation 1: Signal's frequency from the Hertz scale to the Mel-Scale

frequencies that are lesser in Hz have a bigger distance among them Mels what makes the Mel-Scale important in Machine Learning software to audio of imitating humans' insight of sound [27].

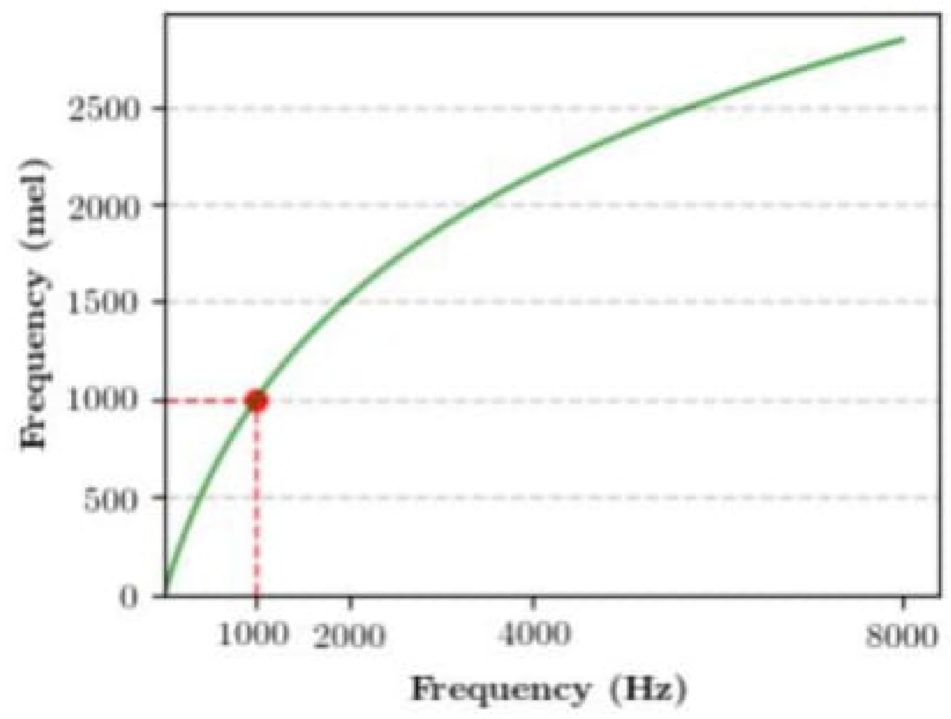


Figure 3 Mel Scale

2.3 The Spectrogram

Spectrograms represents data of frequency domain delivers a means visual data interpretation that represent signal frequency over time [28].

2.4 The Mel Spectrogram

The Mel-spectrogram is one of the efficient methods for audio processing to denote the signal strength of audio over time in order to extract useful features to identify the audio [27]. The Mel spectrum contains a short-time Fourier transform (STFT) for each frame of the spectrum (energy/amplitude spectrum), from the linear frequency scale to the logarithmic Mel-scale [27].

2.5 Convolutional Neural Network

Convolutional Neural Networks (CNNs or ConvNets) in Computer Vision (CV) is currently an extreme fast developing technology that use Artificial Neural Networks (ANN) [29] concept in image processing in Machine Learning (ML) and DL [30]. CNNs are part of a wider family of methods known as DL, a common group of neural networks designed to effectively process image data [31]. In the preceding decade, several enhancements to CNNs, from input exemplification, total of layers, types of pooling, optimization approaches, and applications for a diverse jobs, have been dynamic research issues [32]. Improvements in CNNs includes convolution operations, convolution layers, architecture design, loss functions, and advanced applications [33]. For grid-structured data is a 2-dimensional image, CNNs function best to extract significant characteristics from the spatial correlation and dependencies data using multi-layer perceptron [34] [35].

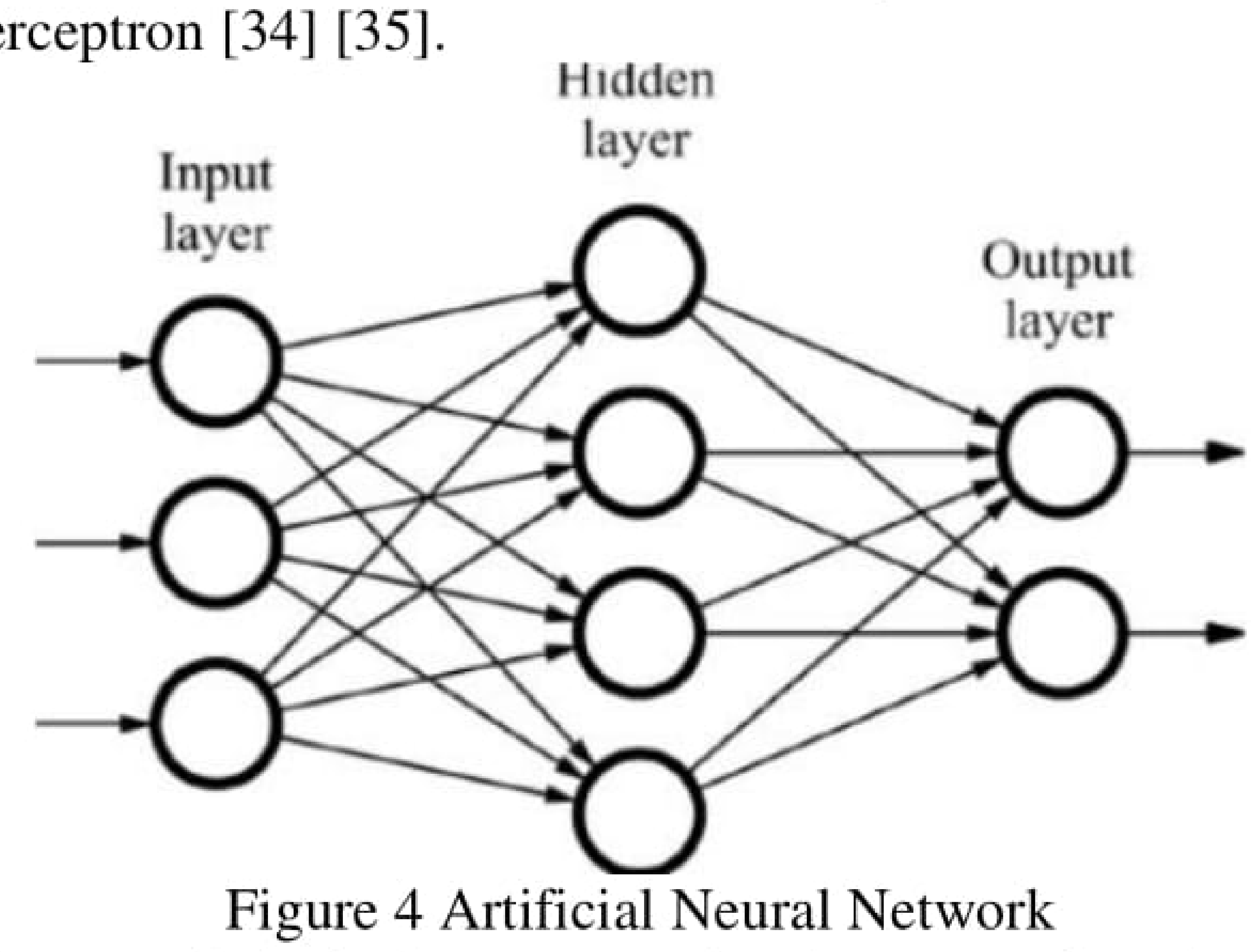


Figure 4 Artificial Neural Network

CNNs are consists of multiple essential elements, including convolution layers, pooling layers, and fully connected layers [36], and are built using a backpropagation algorithm to learn spatial hierarchies of features automatically and adaptively [37].

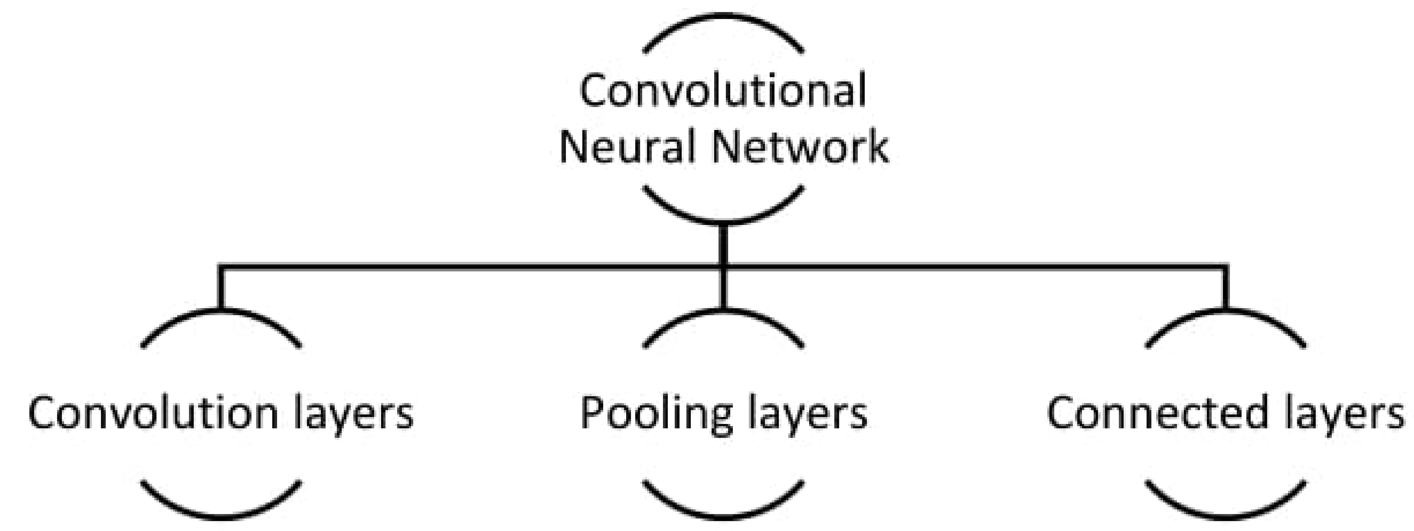


Figure 5 Layers in CNNs

A convolution can be represented as a method from the original image that generates a new image representation. Using the kernel matrix representing the weight, it generates multiple feature maps, and then continues with classification. Mathematically, the convolution is achieve through a kernel of N * N size which sweeps the original image and converts the original data into a different shapes and adding then all the resulting components [38].

3.0 Previous Studies

Several researches have been conducted about heart sound analysis, segmentation which can be used for facilitating the subsequent classification techniques [70]. In this study, we present the most relevant studies within a certain scope classification related to cardiovascular diseases. Computerized heart sound auscultation based on deep learning particularly, convolutional neural network (CNN) [71], has shown that promising results can be obtained in spatial and channel attention [72]. Identifying the normality status of the heart sound recordings (so-called Phonocardiogram - PCG) using multi-classifier of K-Nearest Neighbor [73], Support Vector Machine (SVM) classifiers [74] , frequency domain adaptive line enhancer (FDALE) [75] has been achieved on heart sounds with high accuracy metrics [76] [77]. Classification of normal and pathologic classes using nearest neighbor classifier with Euclidean distance has shown better performance in comparison to Filter banks and Wavelet transform [78]. Heart sound segmentation and classification on a large noisy dataset achieved with average precision of 86.1% using a Markov-switching Autoregressive (MSAR) extended by switching linear dynamic system (SLDS) which significantly outperforms the hidden semi-Markov model (HSMM) in heart sound segmentation [79]. Pre-trained CNN models such as AlexNet, VGG16 [80], and VGG19 spectrogram generation, deep feature extraction, and classification outperformed other existing methods [80]. A competitive method applied using Recurrent Neural Network (RNN) that is based on Long Short-Term Memory (LSTM), Dropout, Dense and Softmax layer [23]. Despite the fact that much research has been conducted in automatic analysis, further work needs to be done to establish reliable methods for identifying and classifying different events in the cardiac cycle [81]. Multiple types of network architectures has been introduced. The first network is VGG 16; which is a very deep ConvNets for largescale image classification with very high accuracy, excellent performance using a conventional [58]. The second architecture is ResNet; which considered a residual learning framework that are substantially deep with layers to learn residual functions with reference to the inputs layer [59]. The third architecture represents MobileNet; one of the streamlined architectures for mobile and embedded vision applications that employs separable convolutions [60]. The forth is InceptionV3; serves as convolutional scale up that aims at utilizing factorized convolutions and batch regularization [61]. Xception is the last network architecture interpreted as a transition between normal convolution and the depth-wise; a single convolutional filter is applied to each input channel in depth-wise convolution, while pointwise convolution is used to generate a linear combination of the depth-wise convolution's output [62].

3.1 Comment about previous Studies

In literature review, there number of gaps and shortcomings in relation to handling of heart sounds as image classification problem. A number of challenges concerning deep models' performance in terms of handling audio as a classification task in frequency domain were addressed. Few researches took closer look on converting the audio data into spectrograms as cardio vascular disease diagnosis. Most studies exclusively focused on image classification in vision field which remain to be limited to a certain extent in in medical field. We are expected to have better results in performance and evaluation metrics.

4. Dataset

Data has been gathered from two sources:

A. From the general public via the iStethoscope Pro iPhone app, provided in dataset A, and
B. From a clinic trial in hospitals using the digital stethoscope DigiScope, provided in dataset B.

The dataset was collected from two sources, A and B:

Table 1 Data set description and details

Data set description	Total	Details
Set A	704	• set_a.csv - Labels and metadata for heart beats collected from the general public via an iPhone app • setatiming.csv - contains gold-standard timing information for the "normal" recordings from Set A.
Set B	1636	• set_b.csv - Labels and metadata for heart beats collected from a clinical trial in hospitals using a digital stethoscope • audio files - Varying lengths, between 1 second and 30 seconds. (Some have been clipped to reduce excessive noise and provide the salient fragment of the sound).

The dataset was split using (80x10x10) into three datasets: Training, Validation, and Testing.
The number of samples of the Training, Validation, and Testing: 1872, 234, 234 respectively.

4.1 Language and tool used

Python, as a main programming language tool in our research, is an ecosystem for deep-learning solutions which bridges the gap between the academic and the industry frameworks. DL approach with Python allows time and focus more on the domain, models, and algorithms.

4.2 Google colab

Google Colaboratory which is referred to as “Google Colab” or “Colab” is used in prototyping machine learning models using hardware options such as GPUs and TPUs. Serverless Jupyter notebooks environment is foundational component for interactive development [82].

4.3 Machine Learning Frameworks: Keras and TensorFlow

Keras is a popular API and learning libraries for neural networks written in Python for procedures of Deep Neural Network and structures of CNN models from basics to advanced techniques [83]. Keras uses the Tensorflow backend to build both shallow and deep models [84].

4.4 Additional Python Libraries

Matplotlib: is a comprehensive plotting library designed for making all kinds of static, animated, and interactive visualizations in Python
Librosa: A library that provides the necessary building blocks necessary for music and audio analysis in python.
Pandas: A powerful open source library use that offers data structures and operations for manipulating numerical tables and time series for Python
Numpy: A comprehensive library for the Python consists of multidimensional array objects and routines to perform mathematical and logical operations.
Seaborn: A python library that provides data visualization and exploratory schemes and statistical graphics.

4.5 Image format

The developed portable networks graphics (PNG) format is versatile and provides a lossless compression scheme [85]. In our work, the input images are down sampled into smaller convolutions using Conv2D and MaxPooling2D layers, which can be seen as a window of the input image.

4.6 Preprocessing

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 6 Issue 1, January - 2022, Pages:7-23

The first step of in preprocessing stage in data mounting google drive. The second step includes extracting the compressed file and gather software versions. The third step involve locating the input folder which contains the sound files. The forth step utilize the data Augmentation and Normalization techniques.

4.7 Data Augmentation

Data augmentation in image classification includes simple techniques, such as cropping, rotating, and flipping input images to generate images of different styles to improve the classifier performance [86].

4.8 Network Architecture

Neural Networks sequentially build high-level features layer by layer that correspond to successive layers [87]. In Keras, a Sequential model is appropriate for a plain stack of layers where inputs have a sequential dependence.

4.8.1 Hear Sounds Analysis Model

The model is based on CNNs which are made up of multiple essential elements, including convolution layers, pooling layers, and fully connected layers, and are built using a backpropagation algorithm to learn spatial dimensions of topographies in an automated and dynamic way.

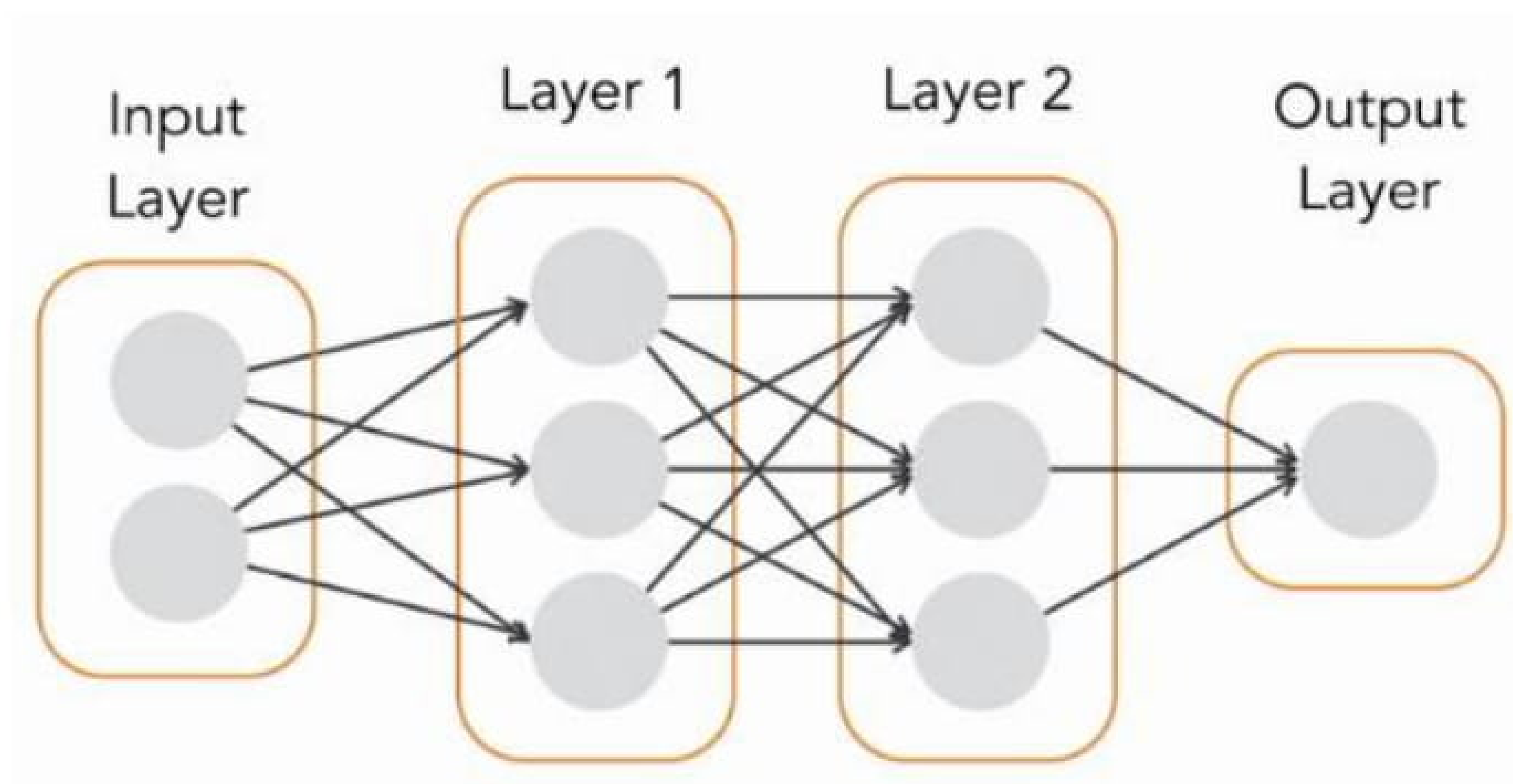


Figure 6 Neural Network Sequential Model using Keras

The workflow of development, training, validating and testing the model was achieved through the steps as shown in following figure:

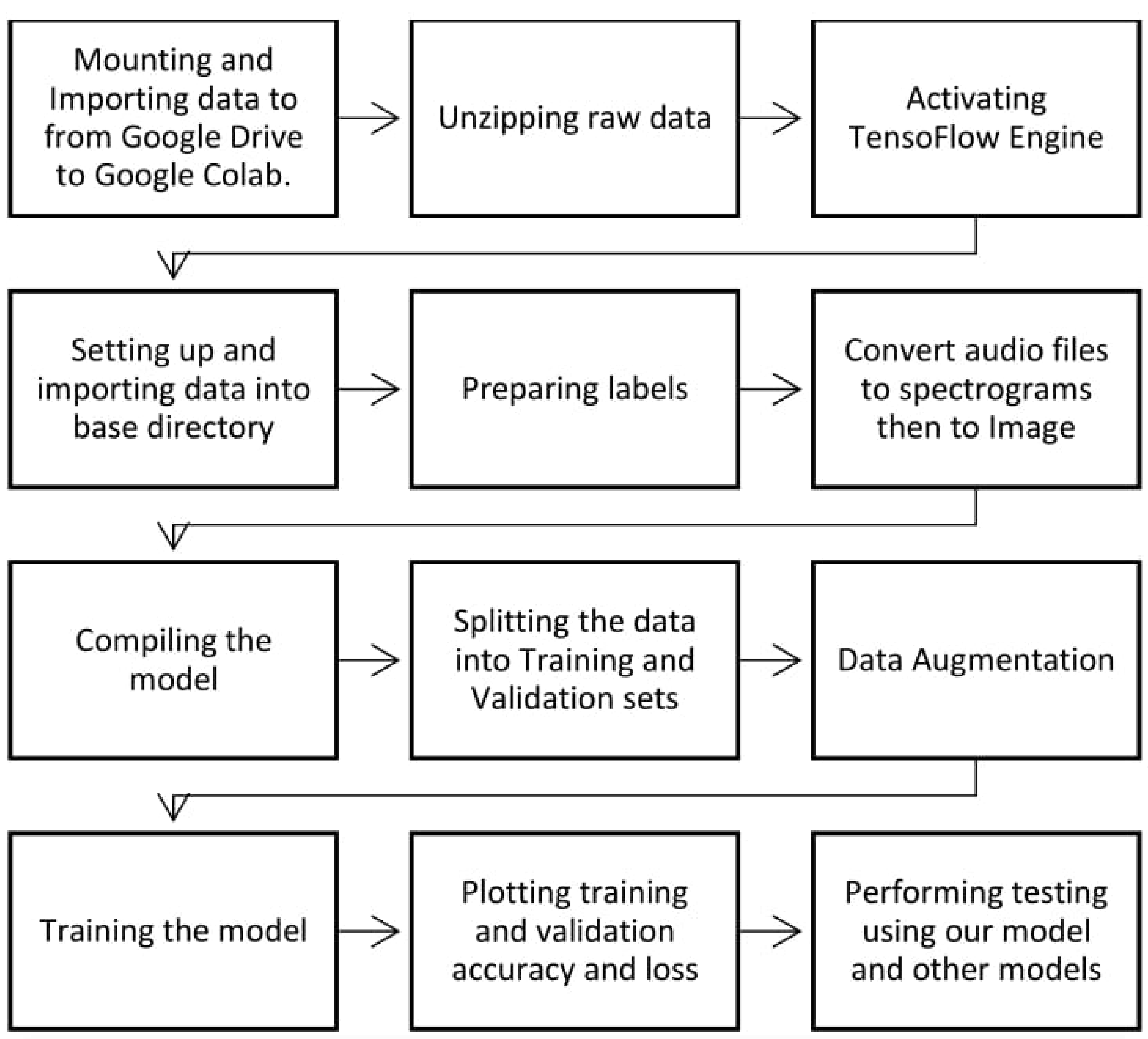


Figure 7 Python Development Workflow using Google Colab

Extracting sound information and frame rates from each sample and save the result as a spectrogram and images.

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 6 Issue 1, January - 2022, Pages:7-23

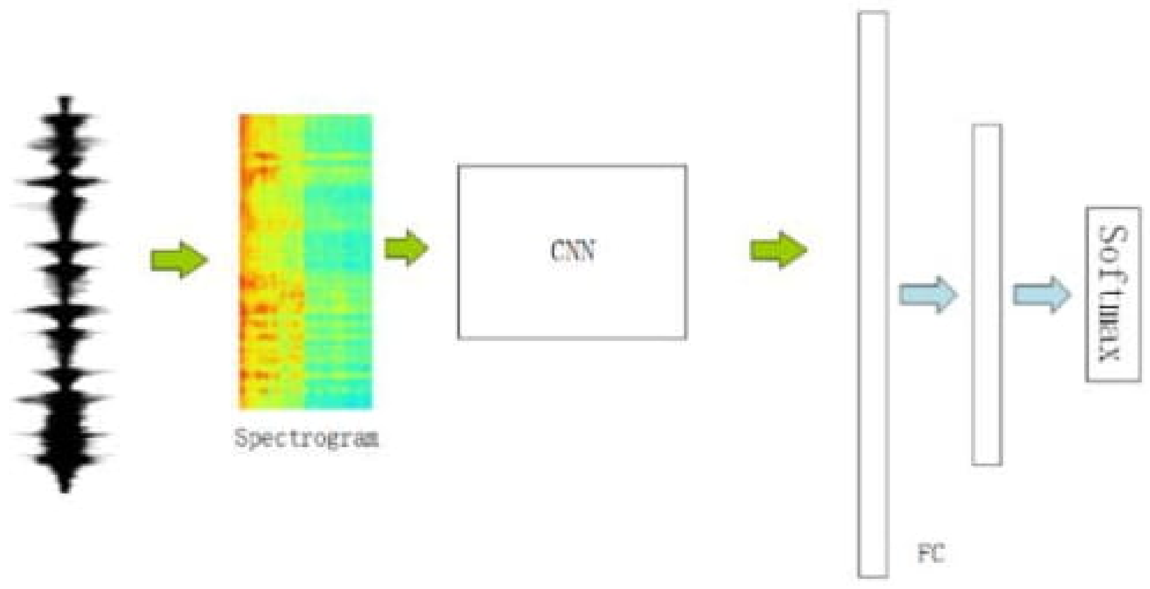


Figure 8 Heart Sounds Classification Model

4.8.2 Data exploration

The audio files are of varying lengths, between 1 second and 30. The majority of info in heart sounds is resides in the low frequency components. On the other hand, noises are contained in the higher frequencies.

As commonly used, applying a low-pass filter is preferred at 195 Hz. In addition, useful information about volume and frequency over time can be obtained using Fast Fourier transforms.

4.8.3 Sound types

4.8.3.1 Normal case

In the Normal category, there are normal, healthy heart sounds. These may contain noise in the final second of the recording as the device is removed from the body, a variety of background noises (from traffic to radios), and occasional random noise corresponding to breathing, or brushing the microphone against clothing or skin.

4.8.3.2 Murmur

Heart murmurs sound as a turbulent fluid noise like roaring in one of two temporal locations: the first could be between “lub” and “dub”, or among “dub” and “lub”.

4.8.3.3 Extrasystole

Extrasystole sounds appear irregularly as the heart become out of rhythm which can be identified though extra or skipped heartbeats: the sounds would be like “lub-lub dub” or a “lub dub-dub”.

4.8.3.4 Artifact

This category has variety of sounds, including feedback squeals and resonances, talking, music and noise without temporal periodicity at frequencies below 195 Hz.

4.8.3.5 Extra Heart Sound

Extra heart sounds can be identified because there is an additional sound, e.g. a “lub-lub dub” or a “lub dub-dub. In addition, extra heart sound cannot be detected by ultrasound.

4.8.4 Audio Length

The lengths in the dataset of the audio files ranges from 1 to 30 seconds. For training purpose we use first 5 seconds of the audio and pad missing length for file smaller than 5 seconds.

4.8.5 Data Handling in Audio domain

Audio data as unstructured data formats has a couple of preprocessing steps proceeding the final analysis. Representing audio data can be accomplished by converting it into frequency domain.

4.8.5.1 **Audio** ***Features for additional analysis***

Converting audio data to MFCs (Mel-Frequency Cepstrums) is an essential step to extract perceptual features in addition to Mel Frequency Cepstral Coefficient (MFCC). Afterwards, machine learning model can be used for class cation purposes or any further analysis.

4.8.5.2 **Sound** ***Feature: MFCC***

MFCC is by far the most successful feature used in Speech Processing. Speech is a non-stationary signal. As such, normal signal processing techniques cannot be directly applied to collectively make up an MFC.

MFCCs are developed following the steps:

1. Produce the Fourier transform of the signal.
2. Using triangular overlapping windows, powers of the spectrum is mapped to Mel scale.
3. For each Mel frequencies, the logs of the powers are taken.
4. Use discrete cosine transform of the list of Mel log powers

The MFCCs are the amplitudes of the resulting spectrum using pre-computed log-power Mel spectrogram. In general, a 39-dimensional feature vector is used which is composed of first 13 MFCCs and their corresponding 13 delta and 13 delta-delta.

4.8.6 Sound Feature: Onset

4.8.6.1 **Onset** ***detector***

By choosing peaks in an onset strength envelope, a basic onset detector may find note onset occurrences. Large-scale hyper-parameter optimization of the dataset supplied was used to choose the peak pick parameters.

4.8.6.2 Onset **backtracks**

Backtrack observed onset occurrences to the energy function's next previous local minimum. This function may be used to reverse the time of detected onsets from a detected peak amplitude to the previous minimum amplitude.

4.8.6.3 **Onset** ***strength***

Compute a spectral flux onset strength envelope. Onset strength at time t is determined by: mean_f max(0, S[f, t] - ref_S[f, t - lag]) where ref_S is S after local max filtering along the frequency axis . S is log-power Mel spectrogram.

The difference in the energy distribution as distinct characteristics the spectrograms of the recordings, which will enable us transform what was originally an audio problem to an image problem.

4.8.7 Convert audio files to spectrograms then to Image (PNG)

Extracting sound information and frame rate from each sample and save the result as a spectrogram and images using pylab library

4.8.8 Preparing the data

The spectrograms are loaded into memory by use the image dataset to generate the datasets, then use Keras image preprocessing layers for image normalization and data augmentation. The batch size is set fairly low for now (32) for all images to fit in memory. The seed is for reproducibility.

4.9 Training and Validation of the of Proposed New Model and other pre-trained models

The following figures shows the training, validation accuracy and loss of each of the proposed model,   all pre-trained models.

Figure 9 Loss curves for training and validation on Proposed New Model

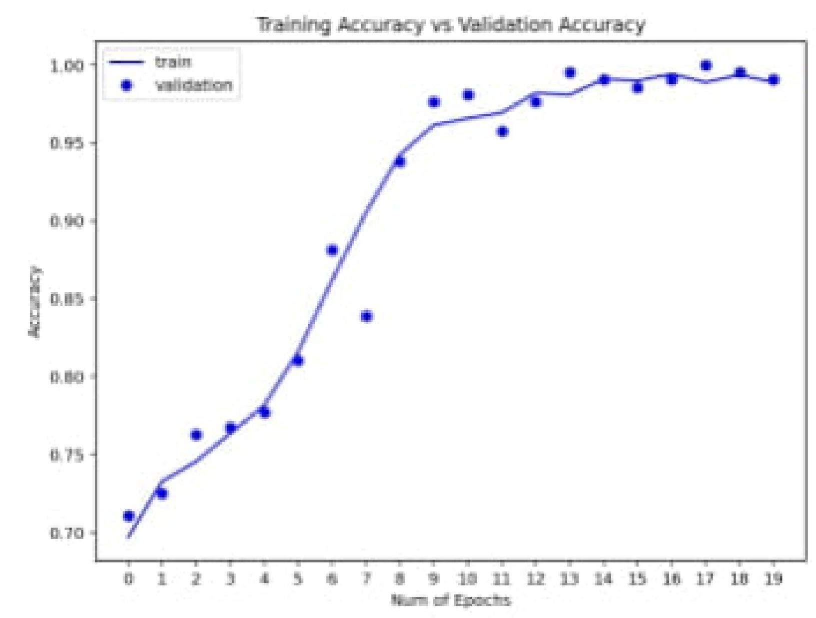


Figure 10 Accuracy curves for training and validation on Proposed New Model

Figure 11 Loss curves for training and validation on VGG16

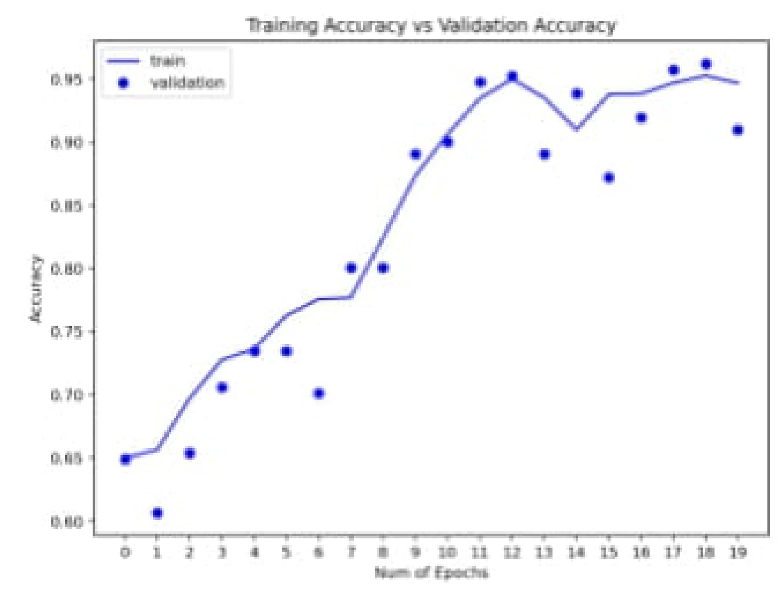


Figure 12 Accuracy curves for training and validation on VGG16

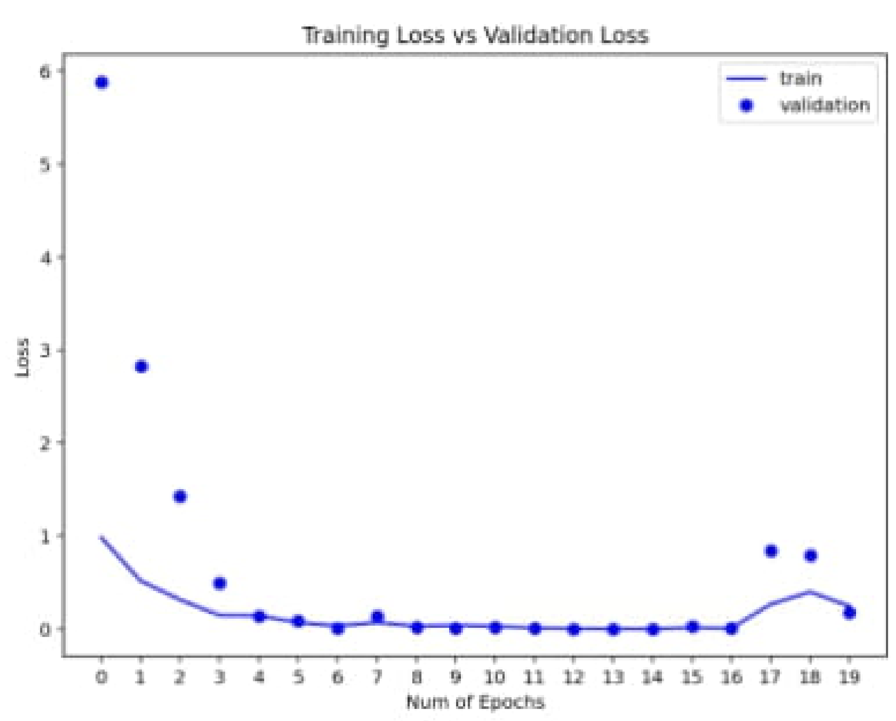


Figure 13 Loss curves for training and validation on ResNet

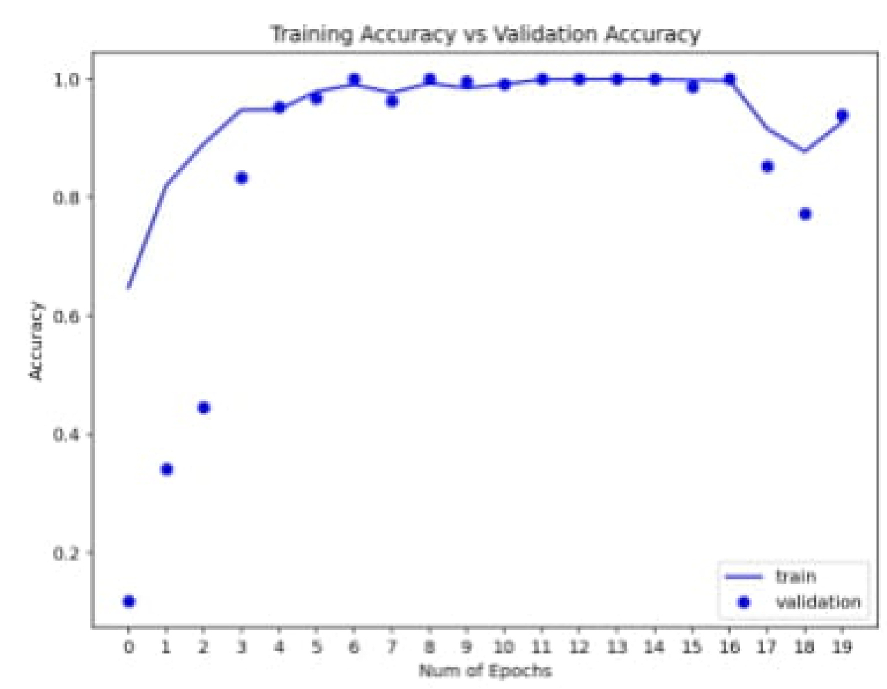


Figure 14 Accuracy curves for training and validation on ResNet

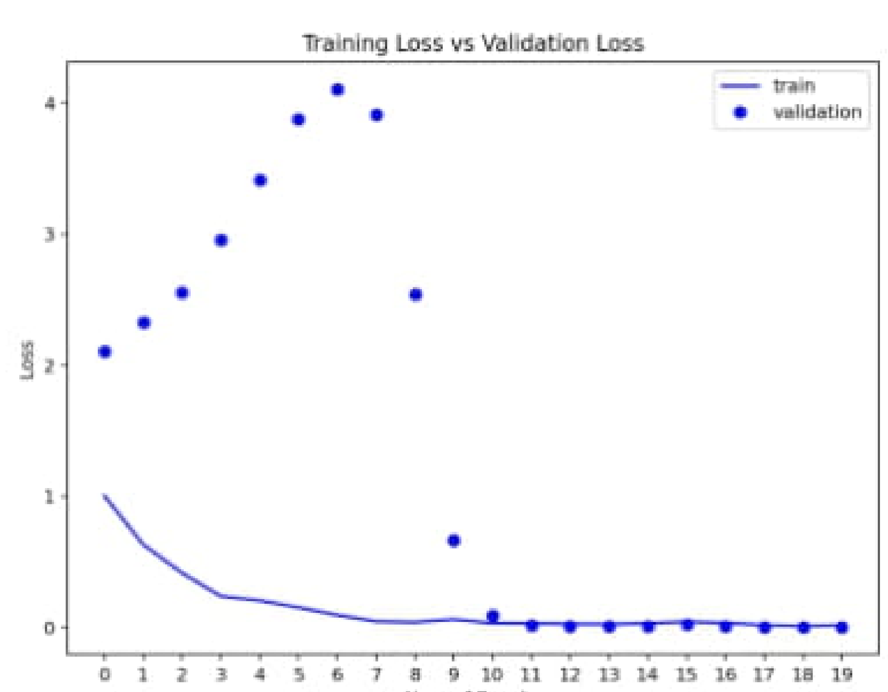


Figure 15 Loss curves for training and validation on MobileNet

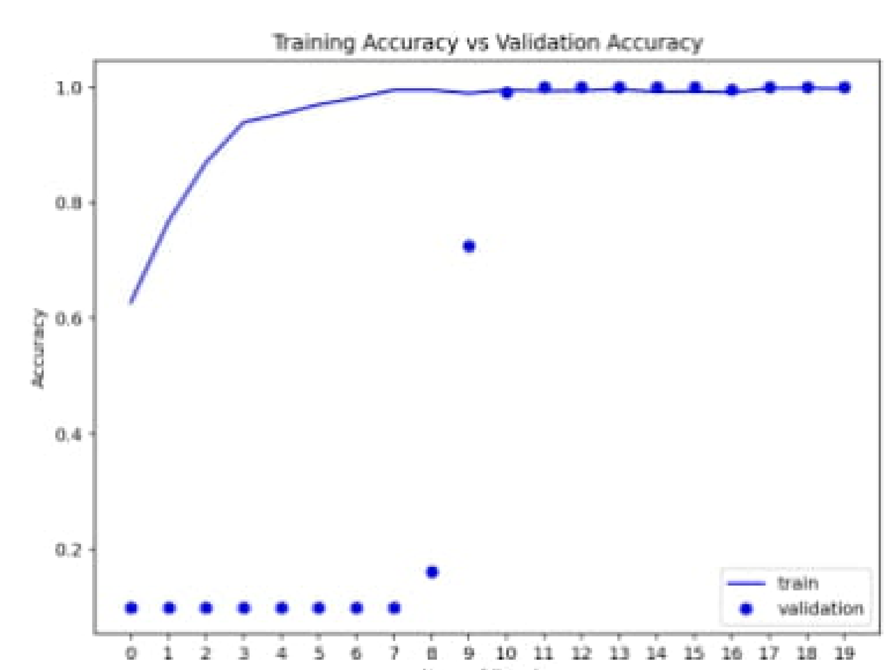


Figure 16 Accuracy curves for training and validation on MobileNet

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 6 Issue 1, January - 2022, Pages:7-23

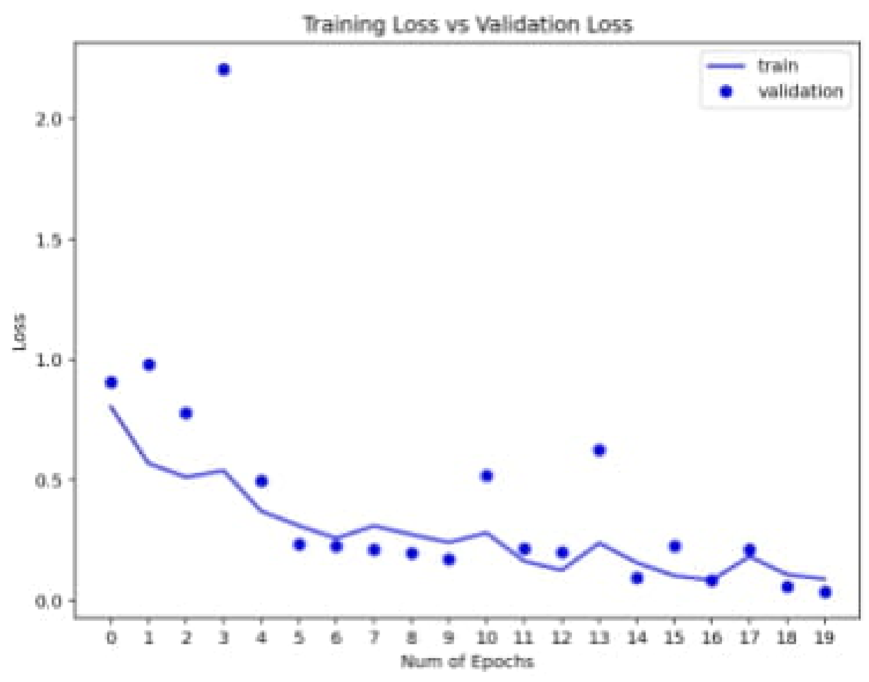


Figure 17 Loss curves for training and validation on InceptionV3 Model

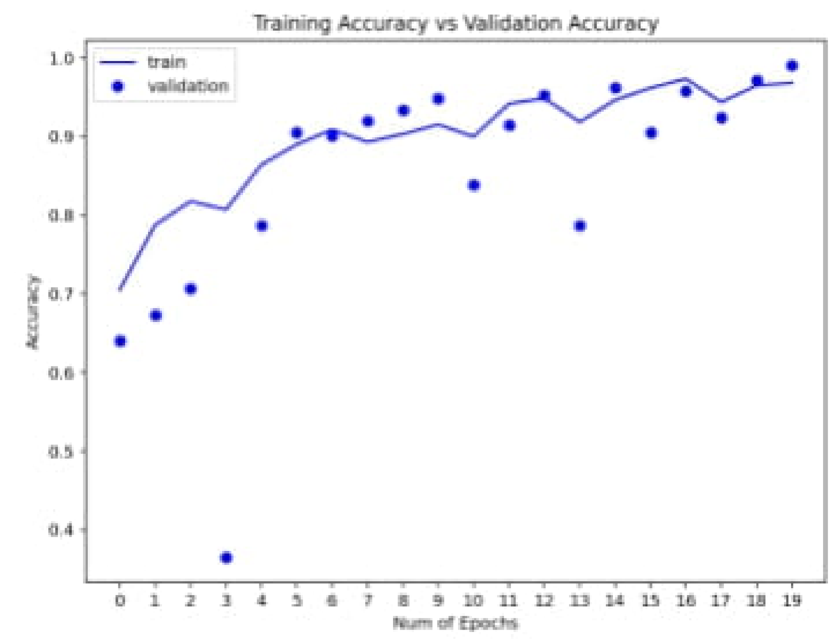


Figure 18 Accuracy curves for training and validation on InceptionV3 Model

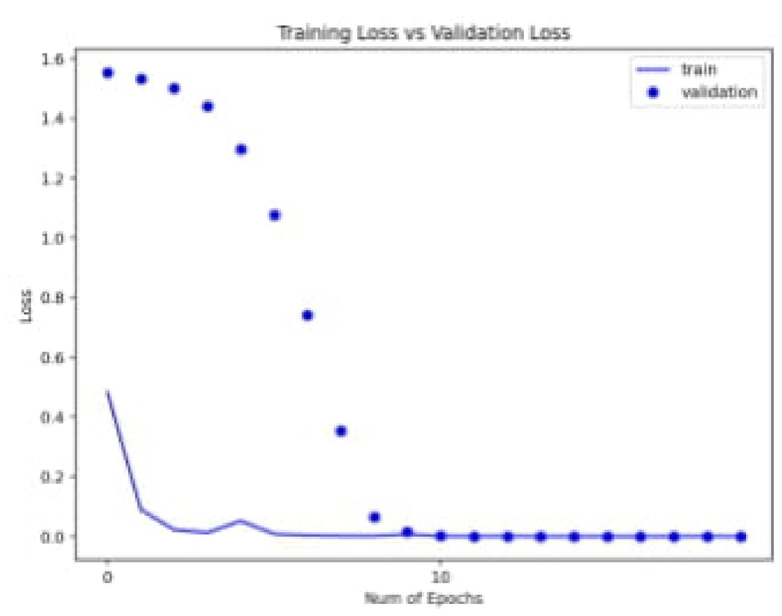


Figure 19 Loss curves for training and validation on Xception Model

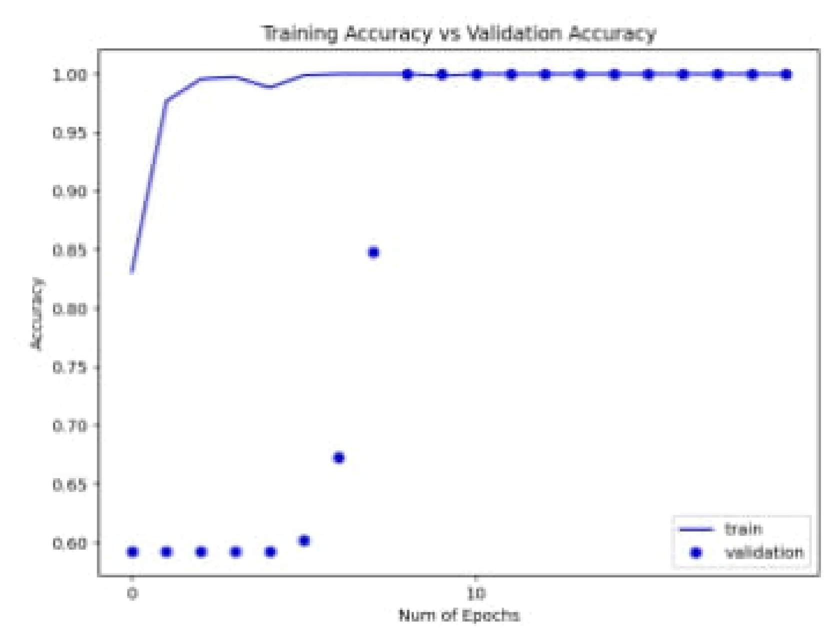


Figure 20 Accuracy curves for training and validation on Xception Model

5. Evaluation of each all models

The dataset was split using (80x10x10) into three datasets: Training, Validation, and Testing.
The number of samples of the Training, Validation, and Testing: 1872, 234, 234 respectively.

In the following figures: Figure 21 shows the accuracy for all models. Figure 22 shows the loss for all models, Figure 23 shows the precision for all models, Figure 24 presents the Recall for all models. Figure 25 presents the F1-Score for all models, Figure 26 presents the Training and Testing Times needed for all models.

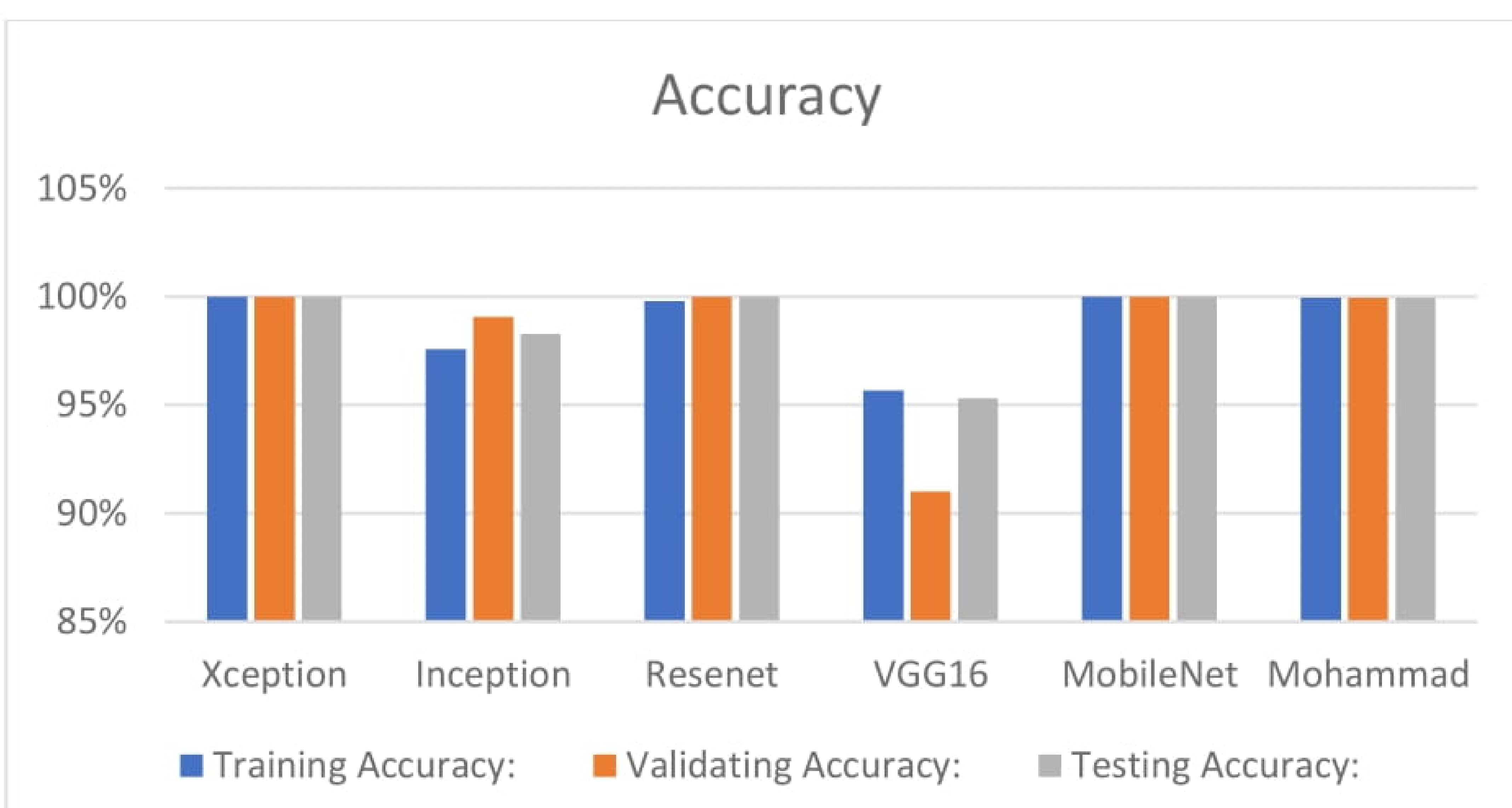


Figure 21 Accuracy for all models

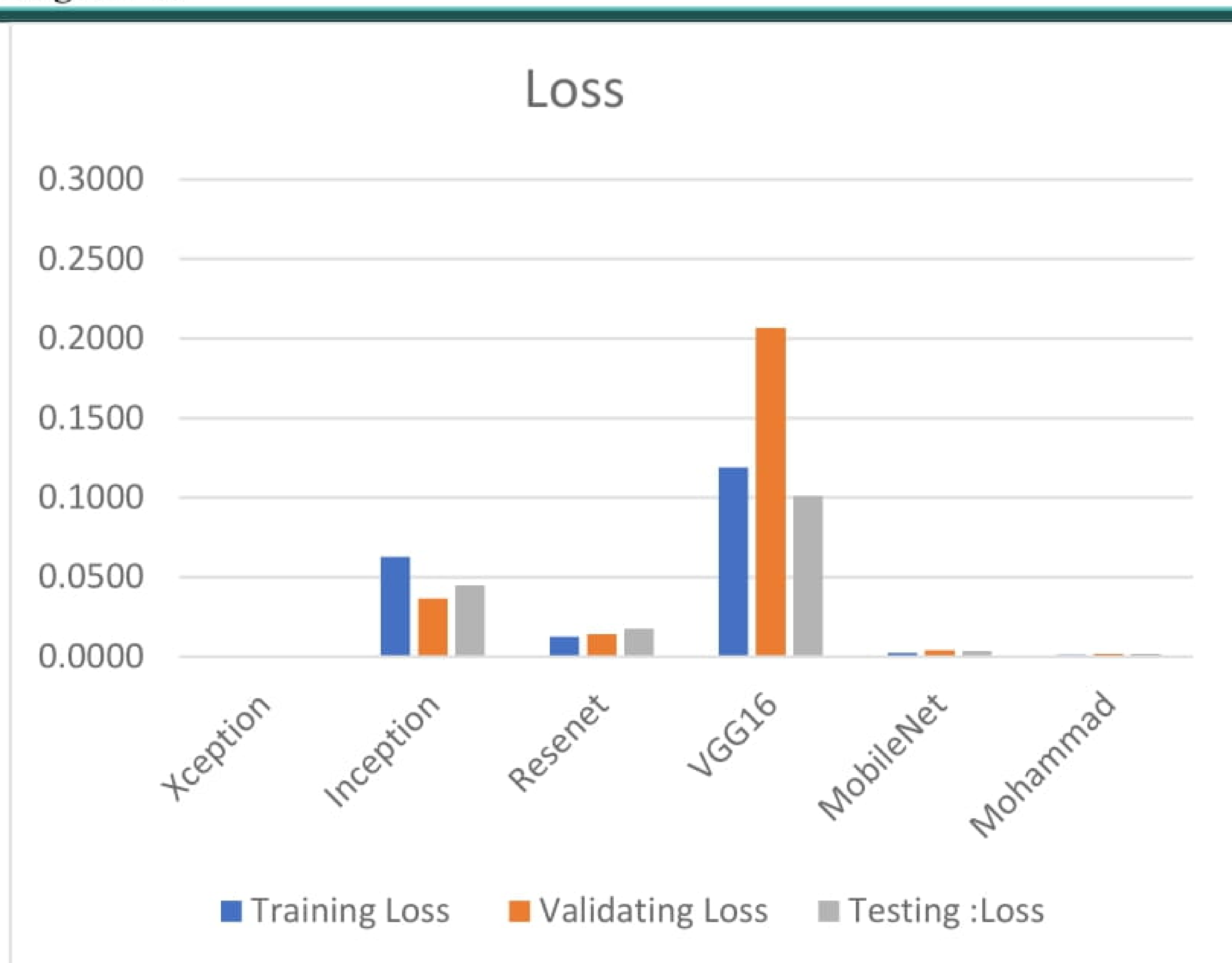


Figure 22 loss for all models

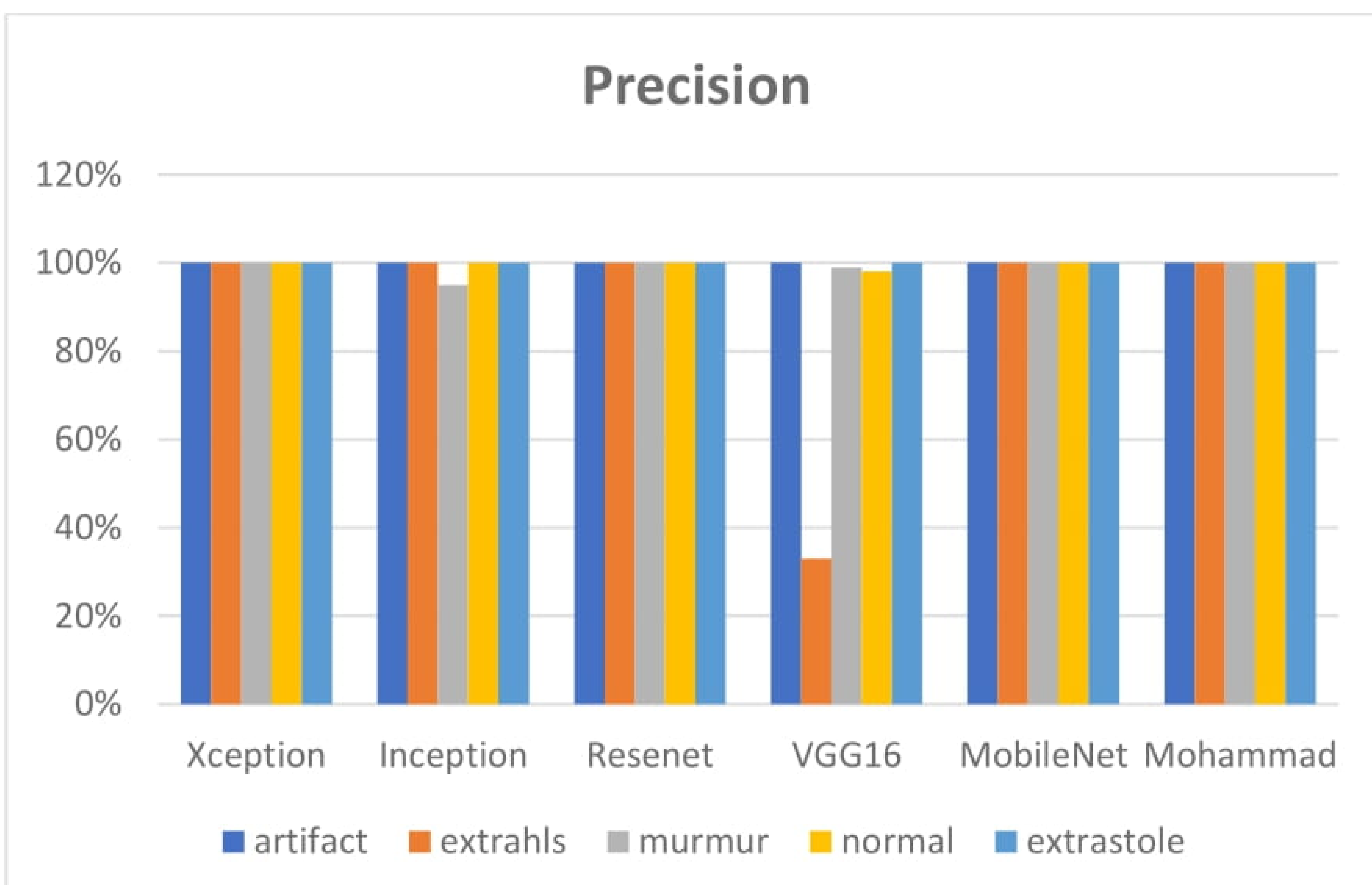


Figure 23 Precision for all models

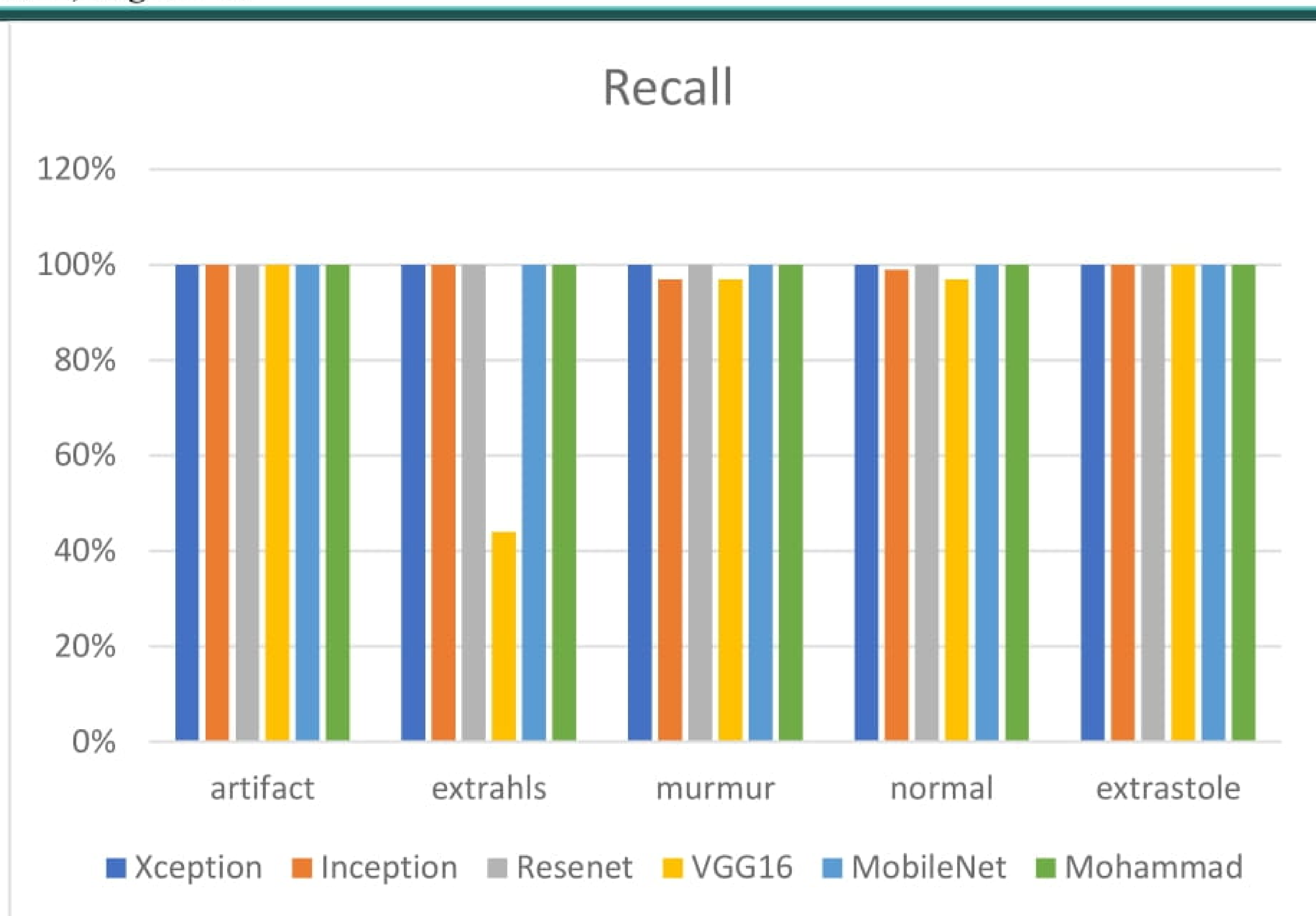


Figure 24 Recall for all models

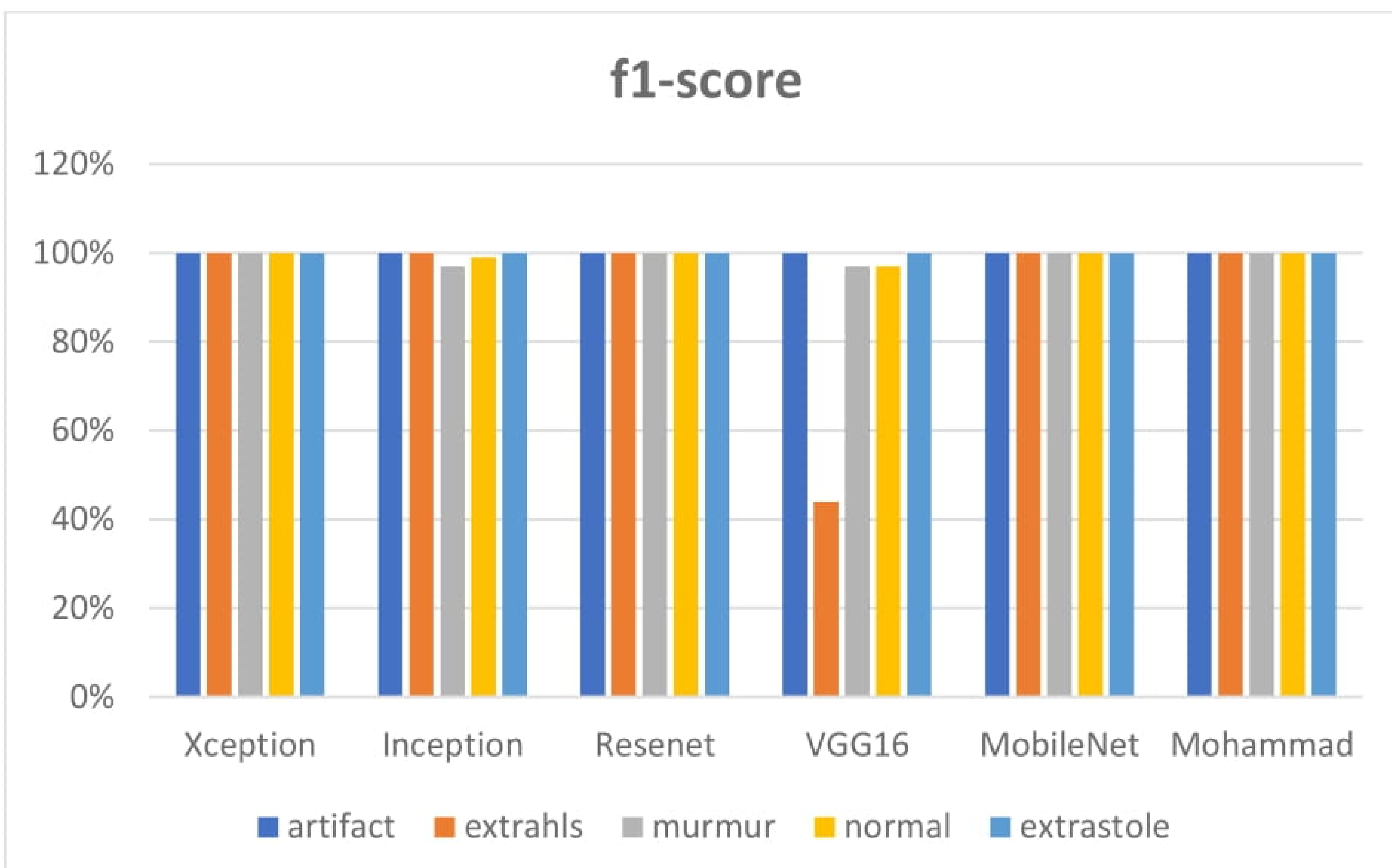


Figure 25 F1-Score for all models

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 6 Issue 1, January - 2022, Pages:7-23

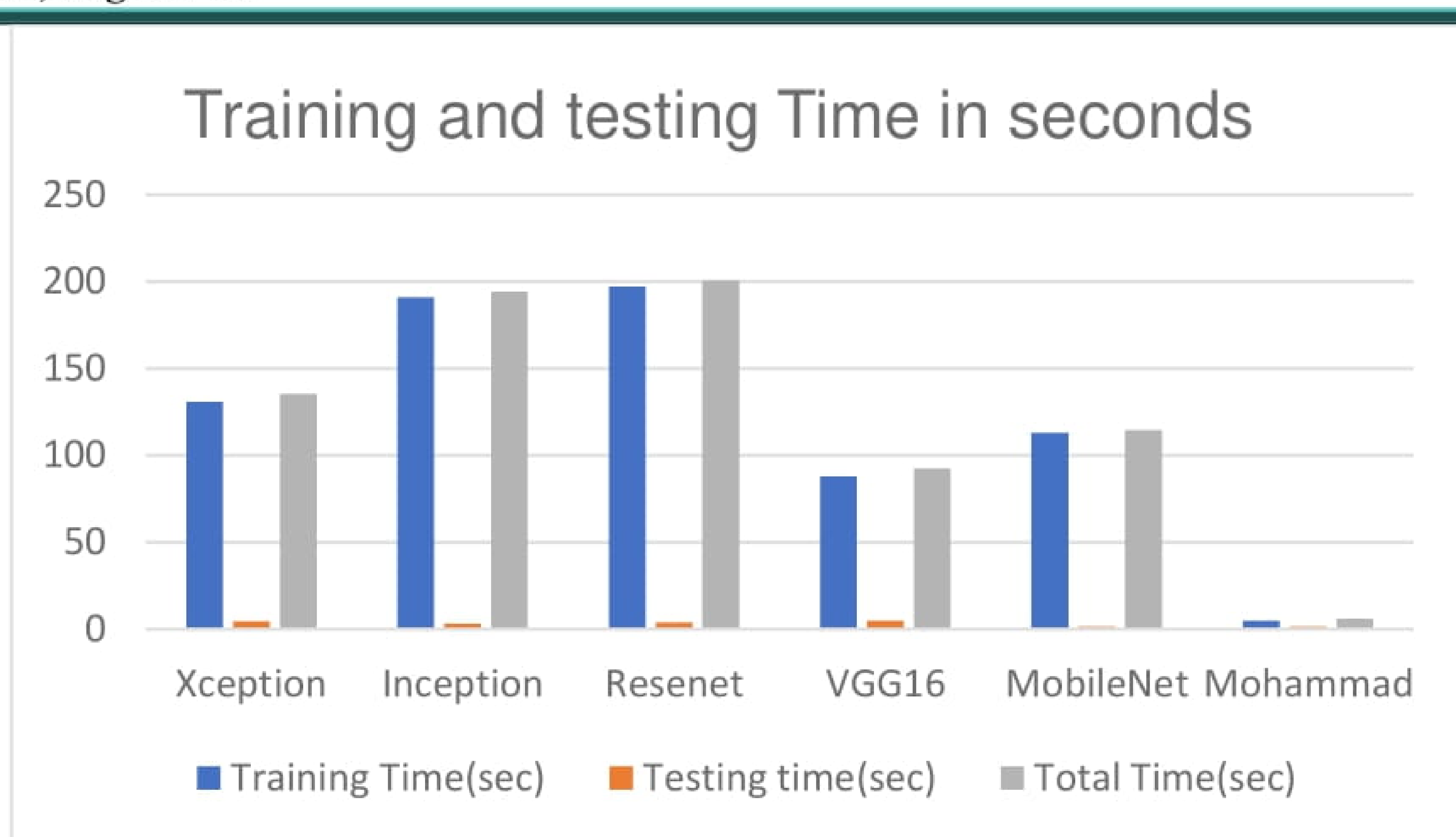


Figure 26 Training and Testing Times for all models

5.2 Result and Discussion of all models

We demonstrate the results of our proposed deep learning model that showed the most robust results in testing and training times which is significantly less in comparison with the other models. In addition, the results go along or beyond achievements previous models in evaluations metric including; Accuracy, Loss, Precision, Recall, and F1-Score. From the results included in classification report and confusion matrices, our finding delivers reliable model that could be used in applications that is concerned with time limitations. The boundaries of our research obviously rely on the publicly used frameworks and cloud platforms that provide sophisticated software and hardware tools for the end users. Using other available tools may alter or improve ranges and aspects of some findings. As discussed, our findings support the notion that further research could be conducted for any shortcomings mentioned in our results.

6. Conclusion

We have introduced a new classification model using Deep Learning (DL) techniques for audio classification using spectrogram. Our method is aimed towards heard sound classification related to cardio vascular disease problems in a way to convert audio data to images which sent to the deep learning model for further analysis. Our method utilizes the advantages of Mel-Frequency Cepstrums (MFC) to extract perceptual features Mel Frequency Cepstral Coefficient (MFCC). The results that have been obtained using our model shows best result in terms of evaluation metric including Accuracy, Loss, Precision, Recall, and F1-Score. Our model has been compared with other models including VGG16, ResNet, MobileNet, Iception V3 and Xception and shown improved results and a reduced amount of training and testing times.

7. Future Work

Our future work will follow several consistent paths which involves enhancing the performance of the model and implementing a real-time system that use portable small-computing and integrated devices and hardware sensors for heart sound detection, acquisition and classification.

REFERENCES

1. Abu Nada, A. M., et al. (2020). "Age and Gender Prediction and Validation through Single User Images Using CNN." International Journal of Academic Engineering Research (IJAER) 4(8): 21-24.
2. Abu Nada, A. M., et al. (2020). "Arabic Text Summarization Using AraBERT Model Using Extractive Text Summarization Approach." International Journal of Academic Information Systems Research (IJAISR) 4(8): 6-9.
3. Abu-Saqer, M. M., et al. (2020). "Type of Grapefruit Classification Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 4(1): 1-5.
4. Afana, M., et al. (2018). "Artificial Neural Network for Forecasting Car Mileage per Gallon in the City." International Journal of Advanced Science and Technology 124: 51-59.
5. Al Barsh, Y. I., et al. (2020). "MPG Prediction Using Artificial Neural Network." International Journal of Academic Information Systems Research (IJAISR) 4(11): 7-16.
6. Alajrami, E., et al. (2019). "Blood Donation Prediction using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 3(10): 1-7.
7. Alajrami, E., et al. (2020). "Handwritten Signature Verification using Deep Learning." International Journal of Academic Multidisciplinary Research (IJAMR) 3(12): 39-44.
8. Al-Araj, R. S. A., et al. (2020). "Classification of Animal Species Using Neural Network." International Journal of Academic Engineering Research (IJAER) 4(10): 23-31.
9. Al-Atrash, Y. E., et al. (2020). "Modeling Cognitive Development of the Balance Scale Task Using ANN." International Journal of Academic Information Systems Research (IJAISR) 4(9): 74-81.
10. Alghoul, A., et al. (2018). "Email Classification Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 8-14.
11. Al-Kahlout, M. M., et al. (2020). "Neural Network Approach to Predict Forest Fires using Meteorological Data." International Journal of Academic Engineering Research (IJAER) 4(9): 68-72.
12. Alkronz, E. S., et al. (2019). "Prediction of Whether Mushroom is Edible or Poisonous Using Back-propagation Neural Network." International Journal of Academic and Applied Research (IJAAR) 3(2): 1-8.
13. Al-Madhoun, O. S. E.-D., et al. (2020). "Low Birth Weight Prediction Using JNN." International Journal of Academic Health and Medical Research (IJAHMR) 4(11): 8-14.
14. Al-Massri, R., et al. (2018). "Classification Prediction of SBRCTs Cancers Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 1-7.
15. Al-Mobayed, A. A., et al. (2020). "Artificial Neural Network for Predicting Car Performance Using JNN." International Journal of Engineering and Information Systems (IJEAIS) 4(9): 139-145.
16. Al-Mubayyed, O. M., et al. (2019). "Predicting Overall Car Performance Using Artificial Neural Network." International Journal of Academic and Applied Research (IJAAR) 3(1): 1-5.
17. Alshawwa, I. A., et al. (2020). "Analyzing Types of Cherry Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 4(1): 1-5.
18. Al-Shawwa, M., et al. (2018). "Predicting Temperature and Humidity in the Surrounding Environment Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 2(9): 1-6.
19. Ashqar, B. A., et al. (2019). "Plant Seedlings Classification Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 3(1): 7-14.
20. Bakr, M. A. H. A., et al. (2020). "Breast Cancer Prediction using JNN." International Journal of Academic Information Systems Research (IJAISR) 4(10): 1-8.
21. Barhoom, A. M., et al. (2019). "Predicting Titanic Survivors using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 3(9): 8-12.
22. Belbeisi, H. Z., et al. (2020). "Effect of Oxygen Consumption of Thylakoid Membranes (Chloroplasts) From Spinach after Inhibition Using JNN." International Journal of Academic Health and Medical Research (IJAHMR) 4(11): 1-7.
23. Dalffa, M. A., et al. (2019). "Tic-Tac-Toe Learning Using Artificial Neural Networks." International Journal of Engineering and Information Systems (IJEAIS) 3(2): 9-19.
24. Dawood, K. J., et al. (2020). "Artificial Neural Network for Mushroom Prediction." International Journal of Academic Information Systems Research (IJAISR) 4(10): 9-17.
25. Dheir, I. M., et al. (2020). "Classifying Nuts Types Using Convolutional Neural Network." International Journal of Academic Information Systems Research (IJAISR) 3(12): 12-18.
26. El-Khatib, M. J., et al. (2019). "Glass Classification Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 3(2): 25-31.
27. 27. Zhou, Q., et al., Cough Recognition Based on Mel-Spectrogram and Convolutional Neural Network. Frontiers in Robotics and AI, 2021. 8(112).
28. 28. French, M. and R. Handy, Spectrograms: Turning Signals into Pictures. Journal of Engineering Technology, 2007. 24: p. 32-35.
29. El-Mahelawi, J. K., et al. (2020). "Tumor Classification Using Artificial Neural Networks." International Journal of Academic Engineering Research (IJAER) 4(11): 8-15.
30. El-Mashharawi, H. Q., et al. (2020). "Grape Type Classification Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 3(12): 41-45.
31. Elzamly, A., et al. (2015). "Classification of Software Risks with Discriminant Analysis Techniques in Software planning Development Process." International Journal of Advanced Science and Technology 81: 35-48.
32. Elzamly, A., et al. (2015). "Predicting Software Analysis Process Risks Using Linear Stepwise Discriminant Analysis: Statistical Methods." Int. J. Adv. Inf. Sci. Technol 38(38): 108-115.
33. Elzamly, A., et al. (2017). "Predicting Critical Cloud Computing Security Issues using Artificial Neural Network (ANNs) Algorithms in Banking Organizations." International Journal of Information Technology and Electrical Engineering 6(2): 40-45.
34. Habib, N. S., et al. (2020). "Presence of Amphibian Species Prediction Using Features Obtained from GIS and Satellite Images." International Journal of Academic and Applied Research (IJAAR) 4(11): 13-22.
35. Harz, H. H., et al. (2020). "Artificial Neural Network for Predicting Diabetes Using JNN." International Journal of Academic Engineering Research (IJAER) 4(10): 14-22.
36. Hassanein, R. A. A., et al. (2020). "Artificial Neural Network for Predicting Workplace Absenteeism." International Journal of Academic Engineering Research (IJAER) 4(9): 62-67.
37. Heriz, H. H., et al. (2018). "English Alphabet Prediction Using Artificial Neural Networks." International Journal of Academic Pedagogical Research (IJAPR) 2(11): 8-14.
38. Jaber, A. S., et al. (2020). "Evolving Efficient Classification Patterns in Lymphography Using EasyNN." International Journal of Academic Information Systems Research (IJAISR) 4(9): 66-73.
39. Kashf, D. W. A., et al. (2018). "Predicting DNA Lung Cancer using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 2(10): 6-13.
40. Khalil, A. J., et al. (2019). "Energy Efficiency Predicting using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 3(9): 1-8.
41. Kweik, O. M. A., et al. (2020). "Artificial Neural Network for Lung Cancer Detection." International Journal of Academic Engineering Research (IJAER) 4(11): 1-7.
42. Maghari, A. M., et al. (2020). "Books' Rating Prediction Using Just Neural Network." International Journal of Engineering and Information Systems (IJEAIS) 4(10): 17-22.
43. Mettleq, A. S. A., et al. (2020). "Mango Classification Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 3(12): 22-29.
44. Metwally, N. F., et al. (2018). "Diagnosis of Hepatitis Virus Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 2(11): 1-7.
45. Mohammed, G. R., et al. (2020). "Predicting the Age of Abalone from Physical Measurements Using Artificial Neural Network." International Journal of Academic and Applied Research (IJAAR) 4(11): 7-12.
46. Musleh, M. M., et al. (2019). "Predicting Liver Patients using Artificial Neural Network." International Journal of Academic Information Systems Research (IJAISR) 3(10): 1-11.
47. Oriban, A. J. A., et al. (2020). "Antibiotic Susceptibility Prediction Using JNN." International Journal of Academic Information Systems Research (IJAISR) 4(11): 1-6.
48. Qwaider, S. R., et al. (2020). "Artificial Neural Network Prediction of the Academic Warning of Students in the Faculty of Engineering and Information Technology in Al-Azhar University-Gaza." International Journal of Academic Information Systems Research (IJAISR) 4(8): 16-22.
49. Sadek, R. M., et al. (2019). "Parkinson's Disease Prediction Using Artificial Neural Network." International Journal of Academic Health and Medical Research (IJAHMR) 3(1): 1-8.
50. Salah, M., et al. (2018). "Predicting Medical Expenses Using Artificial Neural Network." International Journal of Engineering and Information Systems (IJEAIS) 2(20): 11-17.
51. Salman, F. M., et al. (2020). "COVID-19 Detection using Artificial Intelligence." International Journal of Academic Engineering Research (IJAER) 4(3): 18-25.
52. Samra, M. N. A., et al. (2020). "ANN Model for Predicting Protein Localization Sites in Cells." International Journal of Academic and Applied Research (IJAAR) 4(9): 43-50.
53. Shawarib, M. Z. A., et al. (2020). "Breast Cancer Diagnosis and Survival Prediction Using JNN." International Journal of Engineering and Information Systems (IJEAIS) 4(10): 23-30.
54. Zaqout, I., et al. (2015). "Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology." International Journal of Hybrid Information Technology 8(2): 221-228.
55. Saleh, A. et al. (2020). "Brain Tumor Classification Using Deep Learning." 2020 International Conference on Assistive and Rehabilitation Technologies (iCareTech). IEEE, 2020.
56. Almadhoun, H. et al. (2021). "Classification of Alzheimer's Disease Using Traditional Classifiers with Pre-Trained CNN." International Journal of Academic Health and Medical Research (IJAHMR) 5(4):17-21.
57. El-Habil, B. et al. (2022). "Cantaloupe Classification Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 5(12): 7-17.
58. Alkahlout, M. A. et al. (2022). "Classification of Fruits Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 5(12): 56-63.
59. Alfarra, A. H. et al. (2022). "Classification of Pineapple Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 5(12): 37-41.
60. Al-Masawabe, M. M. et al. (2022). "Papaya maturity Classification Using Deep Convolutional Neural Networks." International Journal of Engineering and Information Systems (IJEAIS) 5(12): 60-67.
61. Aldammagh, Z., Abdeljawad, R., & Obaid, T. (2021). Predicting Mobile Banking Adoption: An Integration of TAM and TPB with Trust and Perceived Risk. Financial Internet Quarterly, 17(3), 35-46.
62. Obaid, T. (2021). Predicting Mobile Banking Adoption: An Integration of TAM and TPB with Trust and Perceived Risk. Available at SSRN 3761669.
63. Jouda, H., Abu Jarad, A., Obaid, T., Abu Mdallalah, S., & Awaja, A. (2020). Mobile Banking Adoption: Decomposed Theory of Planned Behavior with Perceived Trust. Available at SSRN 3660403.
64. Obaid, T., Abdaljawad, R., & Mdallalahc, S. A. (2020). Factors Driving E-Learning Adoption In Palestine: An Integration of Technology Acceptance Model And IS Success Model. Available at SSRN 3686490.
65. Obaid, T. F., & Eneizan, B. 84. Chollet, F., keras. 2015.
66. Wiggins, R., et al., Image File Formats: Past, Present, and Future1. Radiographics : a review publication of the Radiological Society of North America, Inc, 2000. 21: p. 789-98.
67. Perez, L. and J. Wang, The Effectiveness of Data Augmentation in Image Classification using Deep Learning. 2017.
68. M. (2016). Transfer of training and post-training on job performance in Middle Eastern countries. Review of Public Administration and Management, 400(3786), 1-11.
69. Obaid, T. F., Zainon, M. S., Eneizan, P. D. B. M., & Wahab, K. A. TRANSFER OF TRAINING AND POST-TRAINING ON JOB PERFORMANCE IN MIDDLE EASTERN COUNTRIES.
70. Chengyu, L., D. Springer, and G. Clifford, Performance of an open-source heart sound segmentation algorithm on eight independent databases. 2017.

71. Li, T., Q. Chunmei, and T. Xiang, Classification of Heart Sounds Based on Convolutional Neural Network. 2017.
72. Xiao, B., et al., Heart sounds classification using a novel 1-D convolutional neural network with extremely low parameter consumption. 2020.
73. Omran, A.-S., et al., Employment of Multi-classifier and Multi-domain Features for PCG Recognition. 2019.
74. Hong, T., et al., PCG Classification Using Multidomain Features and SVM Classifier. 2018.
75. Firuzbakht, F., et al., Abnormal Heart Sound Diagnosis Based on Phonocardiogram Signal Processing. 2018.
76. Omer, D., Classification of Segmented Phonocardiograms by Convolutional Neural Networks. 2019.
77. Malik, S., M. Akram, and I. Siddiqi, Localization and classification of heartbeats using robust adaptive algorithm. 2019.
78. Hamidi, M., H. Ghassemian, and I. Maryam, Classification of heart sound signal using curve fitting and fractal dimension. 2018.
79. Fuad, M.N., et al., A Markov-Switching Model Approach to Heart Sound Segmentation and Classification. 2020.
80. Fatih, D., et al., Towards the classification of heart sounds based on convolutional deep neural network. 2019.
81. Amit Krishna, D., S. Imtiaz, and E. Rodríguez-Villegas, Algorithms for Automatic Analysis and Classification of Heart Sounds–A Systematic Review. 2019.
82. Bisong, E., Google Colaboratory, in Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners, E. Bisong, Editor. 2019, Apress: Berkeley, CA. p. 59-64.
83. Lee, H. and J. Song, Introduction to convolutional neural network using Keras; an understanding from a statistician. Communications for Statistical Applications and Methods, 2019. 26: p. 591-610.
84. Denoyer, L. and P. Gallinari, Deep Sequential Neural Network. 2014.