MUSICAL INSTRUMENT CLSSIFICATION USING DEEP LEARNING APPRROACH

BY

MD. ABUL BASHAR
ID: 161-15-7251

MD. MESBAHUL HOQ
ID: 162-15-7786

SHYMOOM AHMED SOUROV
ID: 162-15-8230

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

MR. ANIRUDDHA RAKSHIT
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

RUBAIYA HAFIZ
Senior Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

OCTOBER 2020

APPROVAL

This Project/internship titled **"Musical Instrument Classification Using Deep Learning Approach"**, submitted by Md. Abul Bashar, ID No: 161-15-7251, Md. Mesbahul Hoq, ID No. 162-15-7786, Shymmom Ahmed Sourov, ID No. 162-15-8230 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 8 October 2020.

BOARD OF EXAMINERS

Dr. Syed Akhter Hossain **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Nazmun Nessa Moon **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Gazi Zahirul Islam **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Saddam Hossain **External Examiner**
Assistant Professor
Department of Computer Science and Engineering
United International University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Mr. Aniruddha Rakshit, Lecturer (Senior Scale),** Department of Computer Science and Engineering**,** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Mr. Aniruddha Rakshit
Lecturer (Senior Scale)
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by:

RUBAIYA HAFIZ
Senior Lecturer
Department of Computer Science and Engineering
Daffodil International University

Submitted by:

MD. ABUL BASHAR
ID: 161-15-7251
Department of CSE
Daffodil International University

SHYMOOM AHMED SOUROV
ID: 162-15-8230
Department of CSE
Daffodil International University

MD. MESBAHUL HOQ
ID: 162-15-7786
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We are really grateful and wish our profound indebtedness to **Mr. Aniruddha Rakshit, Lecturer (Senior Scale)**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of" Machine Learning" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stages have made it possible to complete this project.

We would like to express our heartiest gratitude to **Prof. Dr. Syed Akhter Hossain;** Head, Department of Computer Science and Engineering, for being kind and friendly with us throughout these years and also to other faculty members and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

We have presented a musical instrument classification method in this paper. This Knowledge of the instrumentation of a musical signal could be useful for major audio signal processing problems such as sound source separation and automated music transcription.

We used convolutional neural networks (CNNs), k-Nearest Neighbors (KNN), Support Vector Machine (SVM) to develop the project as it requires less formal statistical training. This project can help people mainly for music data analysis. Many other projects has been prepared earlier.

On this topic but we tried our best to reduce the errors and add more instruments to classify than before.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Music is one of the most popular forms of art that is practiced and listened to by billions of people all over the world. Language of a song can be different but the language of the music created by the instruments are unique. As a part of Music Information Retrieval [5] and music data analysis, musical instrument classification is one of the most crucial tasks in obtaining high-level information regarding a music signal. The primary goal of our project is to classify 17 instruments from music data.

1.2 Motivation

Music is diffusive in our daily life. People from any country, any culture or any caste, there are very less people that don't like music. If you look back you will notice music is one of the amazing arts which is practiced from very old days. Music is found in every known society, past and present, and is considered to be a cultural universal.

Since all people of the world, including the most isolated tribal groups, have a form of music, it may be concluded that music is likely to have been present in the ancestral population prior to the dispersal of humans around the world. People say, music is the medicine of the mind [6]. Whenever we listen to music with an attentive state of our mind, we can feel every single sound from that and it creates a beautiful harmony. We can distinguish between the instruments. From our love to the music we decided to go through this project. It will be very helpful for the people who are working on music data analysis and also the present-day music companies as it can help them on music recommendations for their users.

1.3 Rationale of the Study

- **Instrumentation:** What kind of instruments are present in a music clip and can distinguish the instruments with one another.
- **Music data analysis:** People who are working on music, it will be very much supportive for them. Doesn't matter on what kind of or which country music they are working on.
- **Business purpose:** Nowadays there are several digital music media from where people are listening to music. Like YouTube, Spotify. These media are already using music instrument classifiers in a very advanced way for music recommendation for their user.
- **For Musicians:** There are various kinds of musicians. Some play guitar, some brass instruments, and then some play drums. They can extract a particular instrument sound. Many people using karaoke for vocal practice, where everything except the original vocal tune is missing. This way it can help the musicians.

1.4. Expected Output

People can distinguish between the instruments using an audio file. They can find out which instrument it is.

Musicologists will be benefited most. They used to classify the instruments manually before. Now with the help of machine learning they can make their work done very fast.

When all the instruments and vocals are classified, genre of a music can be predicted easily. Instrument classification is a must thing to do for genre classification.

Improvement will be noticed on digital music media. Students who are studying music will be one step forward on their track.

1.5 Report Layout

Chapter 1: Introduction

In this chapter we discussed the motivation, objective, and the expected outcome of the project. Later followed by the report layout.

Chapter 2: Background

We discuss the background circumstances of our project. We also talk about the related work, comparison to other candidate systems, the scope of the problem and challenges of the project.

Chapter 3: Requirement Specification

This chapter is all about the requirements collection analysis, data collection, proposed methodology and implementation requirements.

Chapter 4: Experimental Results and Discussion

In this chapter we talk about the result of our study and analysis of the result of what we found by experiment.

Chapter 5: Impact on Society, Environment and Sustainability

In this chapter we talk about impact on study, environment and sustainability.

Chapter 6: Summary, Conclusion, Recommendation and Implication for Future Research

In this chapter we summarize our study. We also talk about what we are going to do in future.

CHAPTER 2

BACKGROUND

This chapter is about the theorical explanation of the terminologies used in this research. The chapter also focus on related work abut the topic that have been done previously.

2.1 Preliminaries/Terminologies

The system as a whole is a combination of several technologies e.g. machine learning, Keras, TensorFlow, CNN, SVM, KNN.

2.1.1 Machine Learning

Machine Learning (ML) is an area of computer science that "gives computers the ability to learn without being explicitly programmed" [8]. The Parameter of the formulas is calculated from the data, rather than defined by the programmer. Two most common uses of ML are Classification and Regression [9]. Classification means to categorize different types of data, while Regression means to find a way to describe the data. Basic ML Program will have two stages, fitting and prediction. In the fitting stage, the program will be given a large set (at least thousands) of data. The program will try to adjust its parameters based on some statistical models, in order to make it "fit" the input data best. In the predicting stage, the program will give prediction for a new input based on the parameters it just calculated out. For example, the famous Iris flower dataset [25] contains the measurement of several features of three different species of flowers, such as the length of sepals and petals. A well-defined ML program can learn the pattern behind this feature and give predictions accordingly.

In this project, we will use Jupyter Notebook [10]. It is an open-source web application that contains data cleaning and transformation, machine learning, data visualization etc.

We have applied several models to our dataset for the outcome, like CNN (convolutional neural network, KNN (k-nearest neighbor, SVM (support vector machine), KERAS and TENSORFLOW etc.

2.1.2 CNN

In deep learning, a Convolutional Neural Network [1] is a class of deep neural networks [11], most commonly applied to analyzing visual imagery. Here we introduced a convolutional neural network for music tagging. We used it for local feature extraction and recurrent neural networks for temporal summarization of the extracted features.

We compared CNN with other models that have been used for music tagging while controlling the number of parameters with respect to their performance and training time per sample. Overall, we found that CNNs show a strong performance with respect to the number of parameters and training time, indicating the effectiveness of its hybrid structure in music feature extraction and feature summarization.

Convolutional Neural Networks (CNN) have two types of layer convolutional layers and pooling layers.

Convolutional layers

CNN is patterned to process multidimensional array data in which the convolutional layer takes a stack of feature maps, like the pixels of those color channels, and convolves each feature map with a set of learnable filters to obtain a new stack of output feature maps as input. Output feature map $y(l), l = 1 \ldots \ldots \ldots . L$

$$Y^{(l)} = f(\sum_{k=1}^{k} W^{(k,l)} * X^{(k)} + b^{(l)}) \qquad 2.1.2$$

$x(k)$Stands for the set of K matrices with $k = 1 \dots \dots L$ which comprise the input feature maps, while the $*$operator describes a two-dimensional convolution. Learnable parameters are represented by the weight or kernel matrices $w(k, l)$including the filter coefficients and the bias terms $b(l)$for the respective output feature map$l$. [4]

Pooling layers

Its function is to progressively reduce the spatial size of the representation to reduce the number of parameters and computation in the network, and hence to also control overfitting. The Pooling Layer operates independently on every depth slice of the input and resizes it spatially, using the MAX operation. [22]:

Accepts a volume of size $w1 \times H1 \times D1$

Requires two hyper parameters:

-their spatial extent F

-the stride S

Produces a volume of size $w1 \times H1 \times D1$ where:

$$W2 = \frac{W1-F}{S} + 1$$

$$H2 = \frac{H1-F}{S} + 1$$

$$D2 = D1$$

2.1.3 KNN

K-Nearest Neighbors (KNN) [2] is one of the simplest algorithms used in Machine Learning for regression and classification problems. KNN algorithms use data and classify new data points based on similarity measures. Classification is done by a majority vote to its neighbors. We used KNN for both regression and classification problems. First the data is loaded then k initializes to the selected number of neighbors. The distance calculated between the data, the distance added and the index in an ordered collection, then sort the distances of the ordered collection and the indices from least to leading by the distances. Pick the k entries first from the sorted data, get the labels from the selected entries of k, the mode of k labels will return for classification problems.

2.1.4 SVM:

SVM [3] is a supervised machine learning algorithm which can be used for classification or regression problems. It uses a technique called the kernel trick to transform our data and then based on these transformations it finds an optimal boundary between the possible outputs. Simply put, it does some extremely complex data transformations, then figures out how to separate our data based on the labels or outputs we've defined. Here it finds the optimal linear hyperplane which separates data from different categories with minimum error and maximum margin.

2.1.5 KERAS

Keras [12] is a powerful open source Python library for developing and evaluating deep learning models. It wraps the efficient numerical computation libraries then and TensorFlow and allows us to define and train neural network models.

2.1.7 TENSORFLOW

TensorFlow [13] is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications as we used in neural networks.

2.2 Related Works

Musical Instrument Recognition using CNN and SVM [14], in this paper they mentioned that recognizing musical instruments from any music track is a crucial problem. So, they have proposed a system that will classify the musical instruments of a music track. They used Convolutional Neural Networks (CNN) as it has been extensively researched and has produced highly accurate results for classifying images. Secondly, Support Vector Machines (SVM), as it is used for classification and regression for a long time with a good result. Total nineteen number of instruments can be recognized by their line of action. A high accuracy has been shown while tested by around 6700 training data in their system.

Music and Instrument Classification using Deep Learning Technics [15], they came with an idea on implementation of a multi-class classifier that identifies the instruments in a music stream in this paper. Convolutional neural network is used in the system to preprocess the input audio file and extract the Mel-spectrogram, which is later used for image processing. They focused on three instruments to classify which are piano, drum, flute or other. Around 8000 data were trained in the system and a precision of 70%, a recall of 65% and fi of 64% is achieved for four multiclass classifiers.

Musical Instrument Recognition and Classification Using Time Encoded Signal Processing and Fast Artificial Neural Networks [16]**,** here they described a system for the recognition of musical instruments from isolated notes whereas traditionally, musical instrument recognition is mainly based on frequency domain analysis and shape analysis to extract a set of various features.it is based on time Encoded Signal Processing and Recognition, a time-domain specific feature extraction process. Which is a method to digitally code speech waveforms and then removing all amplitude information by performing infinite clipping.

Classification is performed using Fast Artificial Neural Networks and for validation, they used isolated, constant-pitch notes.

Musical Instrument Classification Using Neural Networks [17]**,** in this paper, a system for automatic classification of musical instrument sounds is introduced. For extraction Mel-Frequency Cepstral Coefficients (MFCCs) are the features used in the system to model the tones as it has been proven to be useful in a broad range of classification applications, such as speech classification, speaker identification etc. In this work, probabilistic neural networks were used for classification for its flexibility and the straightforward design. PNN provides the best matching instrument for each input vector. However, in order to increase the accuracy, multiple input vectors from each sample are needed. Sample database used in experiments consists of 4548 tunes from 19 instruments of the MIS database. The accuracy of 92% was achieved in the best case.

2.3 Comparative Analysis and Summary

In this section, we will explore and analyze different algorithms shown in Table 2.3.1. For the sake of simplicity and a better portrayal different algorithm are grouped by their objective and are described in a tabular manner separately so that the reader can easily get an overall intuition of the models. Table 2.3.1 shows the comparison between different machine learning models trained by audio data.

Table 2.3.1 Comparison between Machine Learning Models

Model	Overview	Evaluation	Reference
CNN	From input data CNN can extract the relevant feature through filters. Filters can learn by CNN automatically	This model achieved 96% Accuracy from Wiley 2020.	[18]
RNN	From the input data RNN can capture the sequential information. Like time series data, text data, and audio data.	This model achieved 77% accuracy from kaggle.	[19]
KNN	This algorithm is used for both classification and regression problems. It produces appropriate output from unlabeled data	This model achieved 98.22% accuracy from IJITEE, May 2020.	[20]
SVM	From input labeled data svm able to categorize new text.	This model achieved 98.22% accuracy from IJITEE, May 2020.	[20]

2.4 Scope of Problem

Our work is to increase accuracy using different types of machine learning models by musical data. It will solve the difficulty of recognizing musical instruments. In this study, we may not gain the expected accuracy but it will be a contribution to MIR [5]

2.5 Challenges

We face many challenges by conducting this research project work in every stage. Problem formulation is the first stage of a research project. Data collection and preprocessing is the most important step of a machine learning project. Then apply a suitable method for the dataset.

CHAPTER 3

Research Methodology

3.1 Research Subject and Instrumentation

Music is such an art that is practiced all over the world. It is mixed with each and every particle of the earthly thing. If we listen to the stream of water of the river, we will find a harmony in that sound. If we listen to the air of the sea, music will also be found there. People who love music, practice music they find everywhere. We proposed our project to enhance the magic of music

3.2 Data Collection Procedure/ Data Utilized

We Collect data from Philharmonia [21]. Phiharmonia has rich sound samples of different types of instruments that is recorded by their members and it is free to use. Dataset consists of 17 different classes of 11497 audio samples. Dataset is divided into Training and Testing data. 8622 samples were used for training and 2875 data were used for testing. We pre-processed data. Mel-spectrograms were used to represent the data and it is the input of our model. To process audio to signal we use sample rate 44100 Hz, an fft size of 2048, hop length of 512.

Figure 3.2.1 shows data processing module

Short Time Fourier Transform

For spectrogram representation we used short time Fourier transform (SFTT), it indicates the time-varying energy across different frequency bands. For a short period of time SFTT is calculated. So called windows of length$N$, to obtain magnitude and phase information per frequency bin K for certain time frame [Muller]

$$X(k) = \sum_{n=0}^{N-1} w(n)x(n)e^{-j2\pi kn/N} \qquad 3.2$$

Herein x denotes a discrete time signal at sample points n sampling ratefs, $w(n)$ describes an N-point temporal window function. The corresponding frequency value for band k is determined by both fs and window length N:

$$f(k) = \frac{k}{N}.f_s \qquad 3.2.1$$

The Mel Scale

On a linear scale, human do not perceive frequencies. In 1937, Volkmann, Newman, Stevens proposed a unit of pitch such that equal distance in pitch sounded quality distant to the lister. This is called Mel scale. [24]

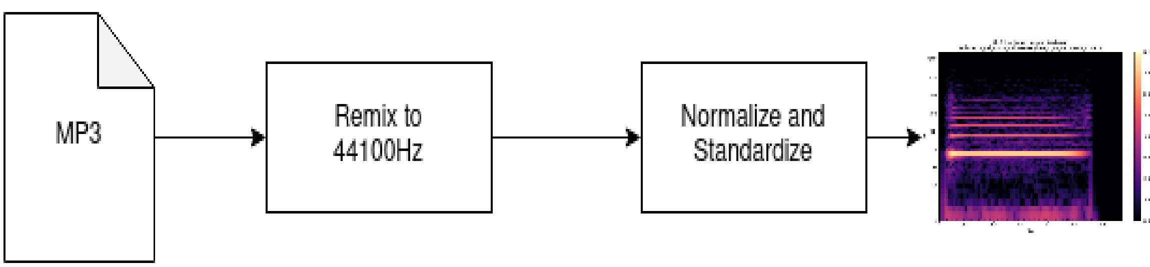

Figure 3.2.1 Data processing module

Figure 3.2.2 show a random spectrogram sample from data set.

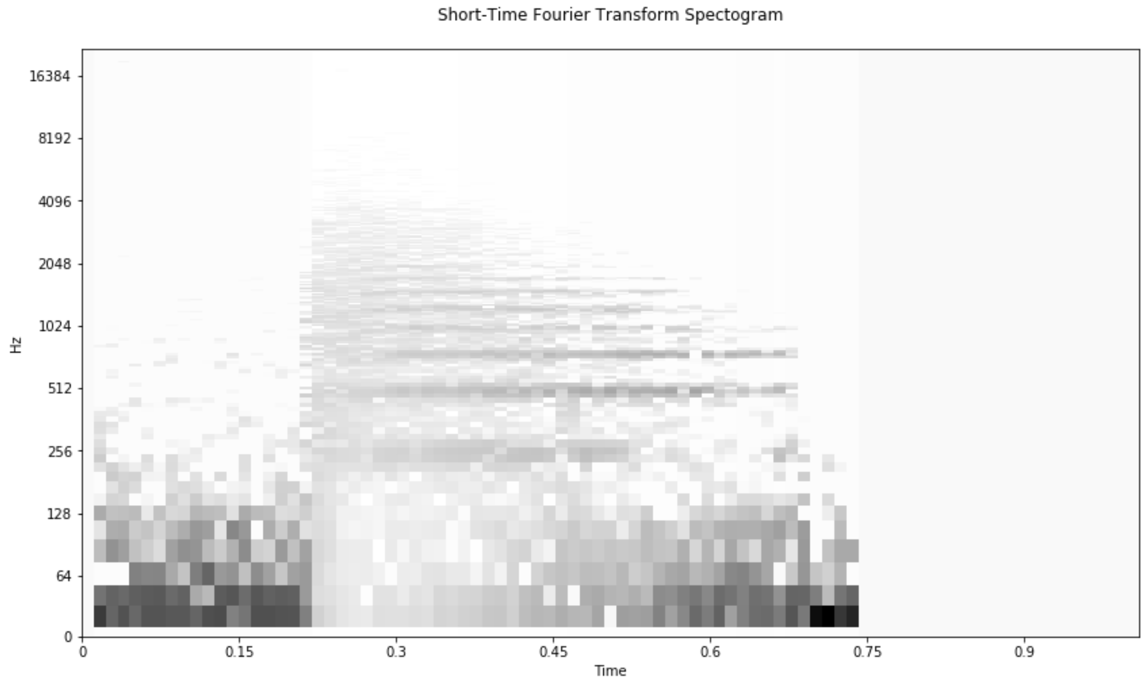

Figure 3.2.2 Random Spectrogram sample

3.3 Dataset Analysis

We worked on 17 different types of class/instrument from Philharmonia [21]. Below table shown the name of instrument with number of samples.

Table 3.3.1 shows dataset of 17 different instrument classes and the number of samples

Table 3.3.1 Dataset analysis

No.	Instrument/ Class Name	Number of Sample
1	saxophone	100
2	mandolin	80

3	banjo	74
4	flute	878
5	trombone	831
6	French horn	652
7	violin	1502
8	clarinet	846
9	tuba	971
10	double bass	851
11	bassoon	720
12	guitar	106
13	cello	888
14	viola	971
15	oboe	596
16	Bass clarinet	943
17	trumpet	485

3.4 Proposed Methodology/ Applied Mechanism

As mentioned previously, first we pre-process our audio data to extract the Mel-spectrogram. Mel-spectrogram that is used as input of the model.

Neural Network Model

In our convolutional neural network model, shown in fig 3.4.1.1 the number of filters is **16**. Maxpool node have a kernel size of **2**. Batch size of 1. Number of epochs **15**. Learning rate of 0.01. We used sigmoid activation to initialize the parameter. We used categorical_crossentropy for loss function and Gradient Descent (“sgd”) as optimizer.

Figure 3.4.1.1 shows the deep learning module that we use in our project.

Max Pooling: To reduce the feature map max pooling is used. Maxpool node have kernel size of (2, 2).

Dropout: Dropout is used to get rid of overfitting. Neural networks tend to overfit the model frequently. In our model we used dropout of 0.4.

Activation Function: Sigmoid activation is used for the input layer in our neural network model. It exists between (0 to 1) that's why we use it to predict probability as output. Softmax function is used for the output layer because we worked on multiclass classification.

Optimizer: Optimizer is used for fine-tuning weight values in a neural network. We use categorical_crossentropy as loss function and Gradient Descent (“sgd”) as optimizer.

Metrics: Different types of metrics are used for the model evaluation. Such as accuracy, recall, precision, threshold etc.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 1023, 85, 16)      160
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 511, 42, 16)       0
_________________________________________________________________
dropout_1 (Dropout)          (None, 511, 42, 16)       0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 509, 40, 16)       2320
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 254, 20, 16)       0
_________________________________________________________________
dropout_2 (Dropout)          (None, 254, 20, 16)       0
_________________________________________________________________
flatten_1 (Flatten)          (None, 81280)             0
_________________________________________________________________
dense_1 (Dense)              (None, 16)                1300496
_________________________________________________________________
dense_2 (Dense)              (None, 6)                 102
=================================================================
Total params: 1,303,078
Trainable params: 1,303,078
Non-trainable params: 0
_________________________________________________________________
```

Figure. Model Summary.

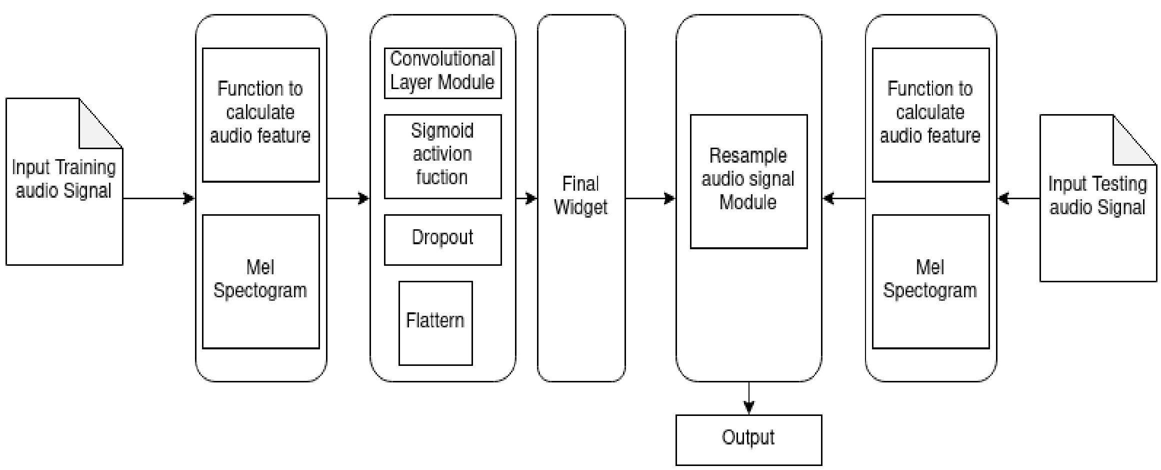

Figure 3.4.1 Architecture of Neural Network proposed work

Classification Model

We used K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) model to train our data set. In Figure 3.4.2.1 we illustrate the module of classification model.

1. **Support Vector Machine (SVM)**

Tuning Parameters: Choosing the right kernel is crucial because it is responsible for poor results from a model. We use kernel of "rbf".

Regularization: To avoid miss classifying each training example Regularization Parameter is used. It tells the SVM optimization how much we want to avoid. In python Regularization Parameter is called "C". We use C of 10.0.

Gamma: We use gamma of 0.10.

2. **K-Nearest Neighbors (kNN)**

We use 1 number of neighbors for the kNN Classifier.

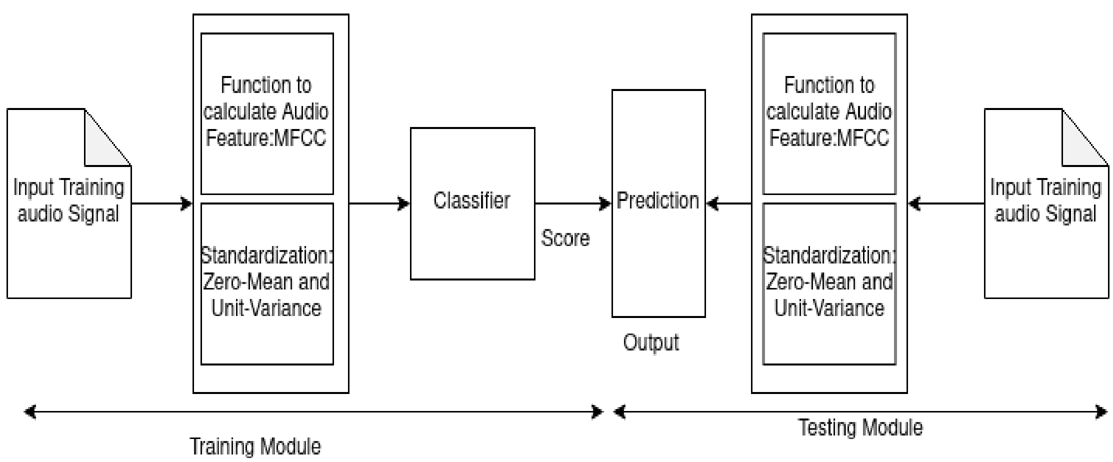

Figure 3.4.2 Architecture of classifier model

3.5 Implementation Requirements

3.5.1 Local System (System Assembling)

- CPU: Intel® Core™ i5-6200U
- Clock Speed: 2600 MHz
- GPU: NVidia GeForce 820M
- RAM: 8 GB
- Video Memory: 2 GB
- L3 Cache: 3 MB
- Operating System: Linux Mint 20.04

3.5.2 Implementation Requirements

- Language: Python (3.7+)
- IDE: Jupyter Notebook, Google Colab

CHAPTER 4

Experimental Result and Discussion

4.1 Experimental Result and Analysis

We trained 11497 audio samples, with input size 1025, 15 epoch and learning rate of 0.01 for classification model. We use 6 classes of 600 music for neural network models. After that we tried multiple hyperparameter to achieve the best result. There is no overfitting.

CNN

We proposed a Convolution Neural Network model that achieved 99% accuracy. Figure 4.1.1 shows training accuracy and validation accuracy over epochs. Figure 4.1.2 shows training and validation loss over epochs.

CNN model evaluation result shows in Table 4.1.4

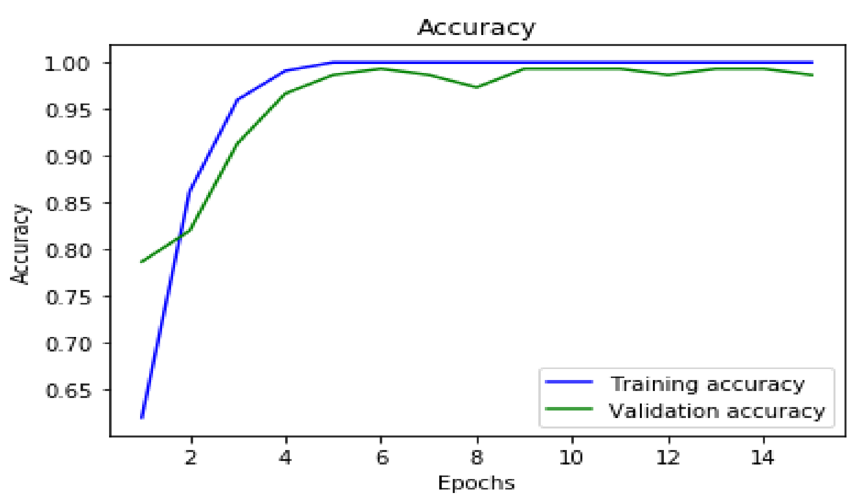

Figure 4.1.1 Training and Validation accuracy over the Epochs

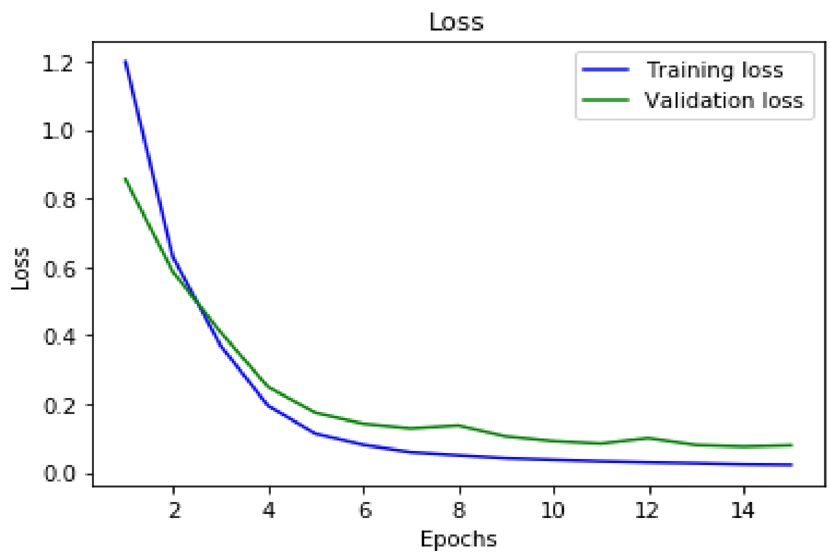

Figure 4.1.2 Training and Validation loss

Figure 4.1.4 shows the confusion matrix of CNN model. This figure is analysis of model using testing data

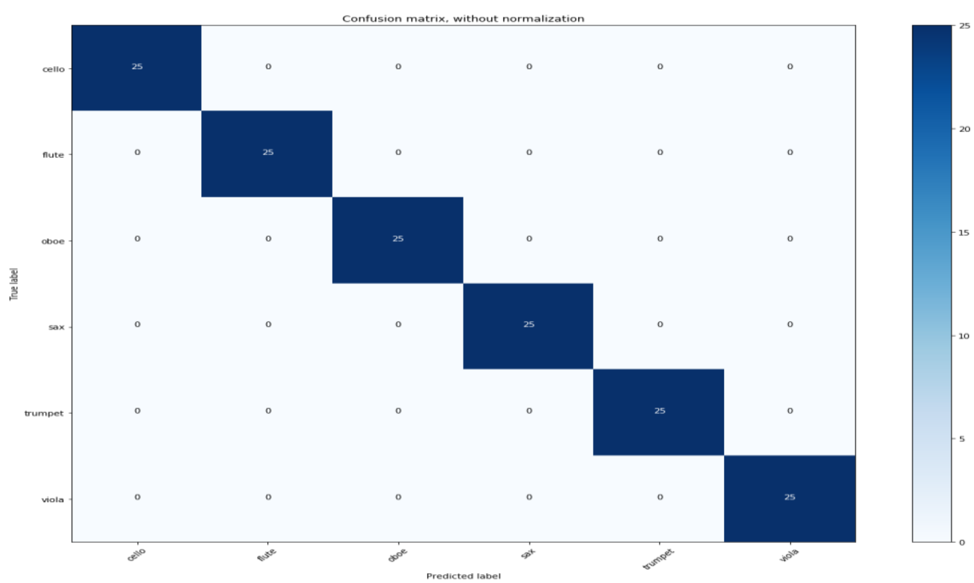

Figure 4.1.3 Confusion matrix of CNN model

Table. 4.1.4 Result for the CNN model for testing data.

Model	**Acc.**	**Recall**	**Precision**	**F1-Score**	**No. of sample**	**Wrong Prediction**	**Loss**
CNN	0.99	0.99	0.99	0.99	150	2	0.0789

SVM

For the testing dataset and classifier model, we computed the confusion matrix of SVM model shows in below figure 4.1.5

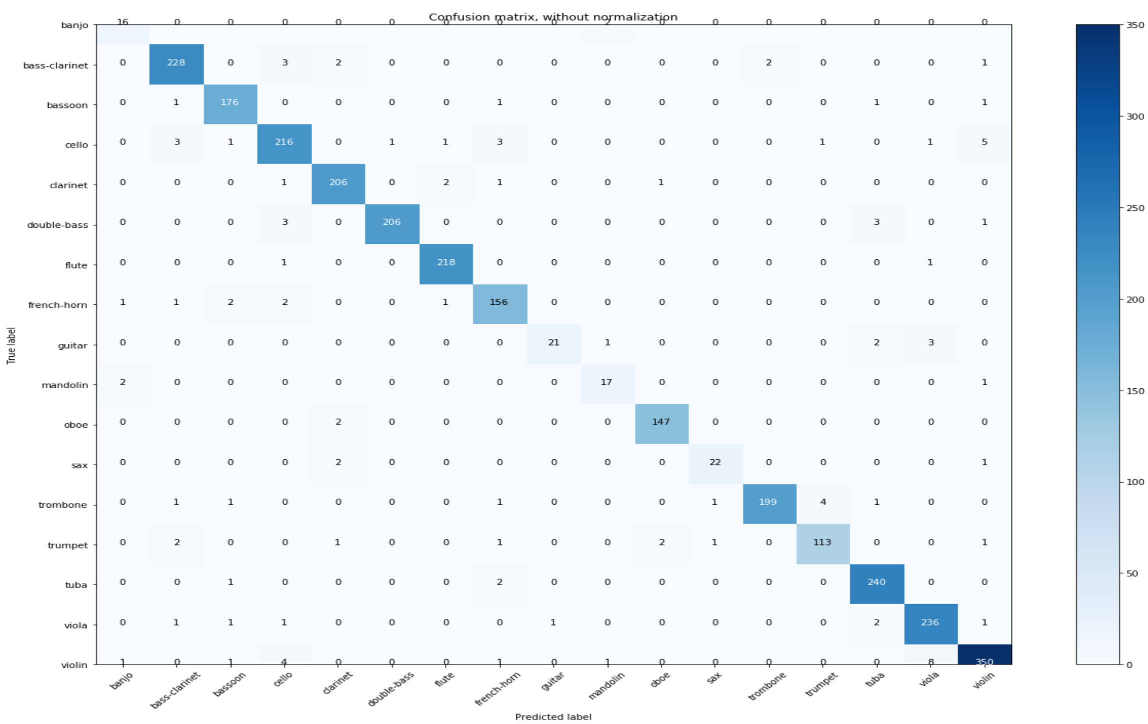

Figure 4.1.5 Confusion matrix of SVM of testing data

KNN

For the testing dataset and classifier model, we computed the confusion matrix of SVM model shows in below figure 4.1.6

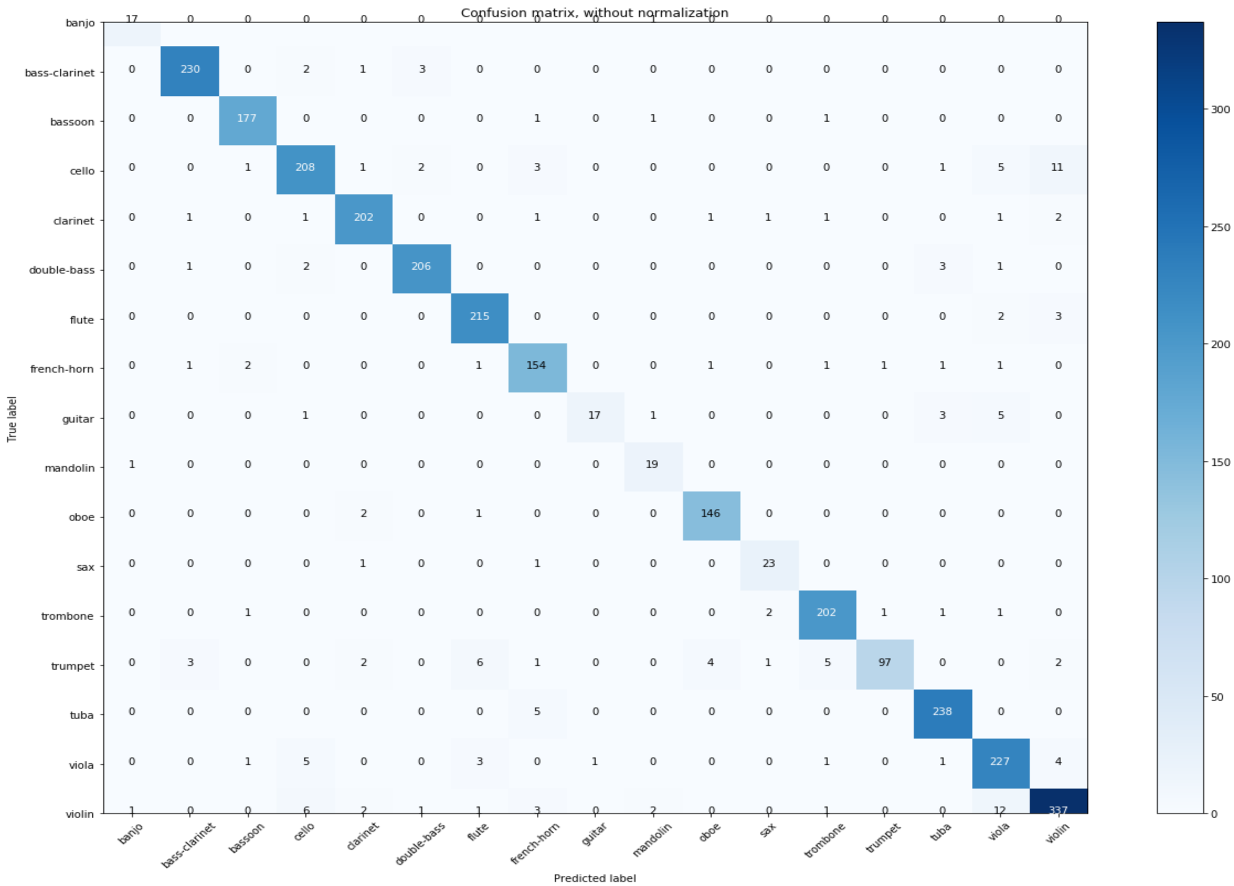

Figure 4.1.6 Confusion matrix of KNN of testing data

Table 4.1.7 shows the comparative analysis between SVM and KNN model for training data. 2875 audio sample were used to testing SVM and KNN model, Table. 4.1.7 Result for the SVM and KNN model for testing data.

Model	Acc.	Recall	Precision	F1-Score	No. of sample	Wrong Prediction
SVM	0.96	0.94	.94	0.93	2875	108
KNN	0.94	0.92	0.93	0.92	2875	160

4.2 Discussion

After all the explanatory analysis of the result given in section 4.1, we can say that our machine learning model is provided a satisfactory result. The hyper tuning of classification models also showed higher accuracy. We have limitations of data; in future we will work on it.

CHAPTER 5

Impact on Society, Environmental and Sustainability

5.1 Impact on Society

As the purpose of our system is to point out the musical instruments from a music clip, and music is such a thing that not only musicians but most human beings love it, so it might have some impact on society. There are different instruments used in different music's like for string instruments, in our local music 'ektara' and 'dotara' are used mostly and they are used in our culture from the very beginning. Whereas in western music guitar and zithers are used in Arabian music. So, when we get to finish our project people will be able to know about the instruments used in their favorite music. Nowadays people are interested in the music of all around the world as social media just made us very close with others. It will be amazing when we will listen to any music of Africa or Middle East and we will know about their music, their instruments. This way people might be attracted to the musical platform more than before and that might be a big change. And surely it will have a big effect on the music industry and who are directly or indirectly attached with the world of music.

5.4 Ethical Aspect

We are developing a project where the system will classify the name of musical instruments from a music clip or music data. We are building this system for the musicians who compose and sing or the people who work in the music industry behind making the beautiful music that gives us pleasure, relaxation during hard times. It is not like it will do everything for them but it will help them a bit in their work. Music is like magic, it makes us happy, it makes us sad, it makes us active, it makes us emotional. We have seen the shyest boy of our batch jumping with the beats during a concert.

With the happiest expression on his face. We are developing this project for those magicians who surprises us with amazing music.

CHAPTER 6

Summary, Conclusion, Recommendation and Implication for Future Research

6.1 Summary of the Study

The title of our proposed project is 'Musical instrument classification'. Which will classify musical instruments from a music clip. For the operation we have fetched data from the internet. Which was processed and modified and then with the 75 percent of data we trained the machine and with the remaining 25 percent we tested our system. After testing the data we found the good accuracy. It was then hyper parameter tuning had reached at the better accuracy. Our system is capable of classifying a total of 17 numbers of instruments from the testing data.

6.2 Conclusions

Since music is a fundamental constituent of human life, our project will help the people around the world, especially those who love music, who have involvement with musical instruments and the industry? We are very much optimistic with our project although it has some shortcomings and needs some development also, we are going to have all of them fixed as in our sustainability plan.

6.3 Implication for Further Study

For the present moment our system can classify a certain category of musical instruments from a music clip of our trained data. We are destined to explore this system for more output.it will classify the genre of the music and we will be able to check it for any music data with any format. It might take a very long time as we are very beginner on this, but this is what we are intended to do.

References

[1] Convolutional Neural Network, available at https://towardsdatascience.com/covolutional-neural-network-cb0883dd6529, last accessed on 06-09-2020 at 3:05 PM

[2] Introduction to k-Nearest Neighbors: A powerful Machine Learning Algorithm, available at https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering,last accessed on 06-09-2020 at 3:05 PM

[3] An Introduction to Support Vector Machines (SVM), available at https://monkeylearn.com/blog/introduction-to-support-vector-machines-svm/, last accessed on 08-09-2020 at 10:00 AM

[4] Fabian Seipel. (2018). Music Instrument Identification using Convolutional Neural Networks. Technische Universit ̈at Berlin

[5] Music Information Retrieval, available at, https://musicinformationretrieval.com/index.html, last accessed on 10-08-2020 at 3:00 PM

[6] John A. Logan Quotes, available at https://www.brainyquote.com/quotes/john_a_logan_209813, last accessed on 12-09-2020 at 10:00 AM

[8] What is machine learning, available at https://theconversation.com/what-is-machine-learning-76759t, last accessed on 08-09-2020 at 12:00 PM

[9] ML | Classification vs Regression, available at https://www.geeksforgeeks.org/ml-classification-vs-regression/, last accessed on 08-09-2020 at 10:20 AM

[10] Jupyter, available at https://jupyter.org/

[11] Deep Neural Network, available at https://machinelearningmastery.com/what-is-deep-learning/, last accessed on 08-09-2020 at 12:07 PM

[12] Keras, available at https://keras.io/

[13] TensorFlow, available at https://www.tensorflow.org/

[14] Prabhjyot Singh, Dnyaneshwar Bachhav, Omkar Joshi, Nita Patil. (2019). Musical Instrument Recognition using CNN and SVM. International Research Journal of Engineering and Technology (IRJET). Volume: 06 Issue: 03 Mar 2019

[15] Lara Haidar-Ahmad. (2019). Music and Instrument Classification using DeepLearning Technics. Stanford University. Cs230

[16] Giorgos Mazarakis, Panagiotis Tzevelekos, Georgios Kouropetroglou. (2006). Musical Instrument Recognition and Classification Using Time Encoded Signal Processing and Fast Artificial Neural Networks. Conference: Advances in Artificial Intelligence, 4th Helenic Conference on AI, SETN 2006, Heraklion, Crete, Greece, May 18-20, 2006

[17] Mustafa Sarimollaoglu, Coskun Bayrak. (2006). Musical Instrument Classification Using Neural Networks. Proceedings of the 5th WSEAS International Conference on Signal Processing, Istanbul, Turkey, May 27-29, 2006 (pp151-154)

[18] Dutta A, Sil D, Chandra A, Palit S. CNN based musical instrument identification using time-frequency localized features.Internet Technology Letters. 2020; e191. https://doi.org/10.1002/itl2.191

[19] RNN musical instrument classification, available at https://www.kaggle.com/yjl450/rnn-musical-instrument-classification, last accessed on 12-08-2020 at 10:00 AM

[20] S. Prabavathy, V. Rathikarani, P. Dhanalakshmi. Classification of Musical Instruments using SVM and KNN. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-7, May 2020

[21] Philharmonia, available at https://philharmonia.co.uk/

[22] Convolutional Neural Networks (CNNs/ConvNets), available at https://cs231n.github.io/convolutional-networks/#pool, last accessed on 02-10-2020 at 11:10 AM

[23] Meinard Muller, Daniel PW Ellis, Anssi Klapuri, and Ga ̈el Richard. Signal processing formusic analysis.IEEE Journal of Selected Topics in Signal Processing, 5(6):1088–1110, 2011.

[24] Understanding the Mel Spectrogram, available at https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53, last accessed on 06-09-2020, 9:05 PM

[25] Iris Flower Dataset, available at https://www.kaggle.com/arshid/iris-flower-dataset, last accessed on 07-08-2020, 10:09 PM

Plagiarism Report

Musical Instrument Classification

ORIGINALITY REPORT

20%	%	%	20%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS