

Music-Source-Separation

BY

Sadiqur Rahman

ID: 162-15-8181

Alam Shahriar

ID: 162-15-8179

AND

Md. Alim-Ul-Alam Masum

ID: 162-15-8227

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

Supervised By

Raja Tariqul Hasan Tusher

Sr. Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

Aniruddha Rakshit

Sr. Lecturer

Department of CSE

Daffodil International University



**DAFFODIL INTERNATIONAL UNIVERSITYDHAKA,
BANGLADESH**

JULY 2020

i

APPROVAL

This Project titled “**Music-Source-Separation**”, submitted by Sadiqur Rahman ID: 162-15-8181, Alam Shahriar ID: 162-15-8179, and Md. Alim-Ul-Masum ID: 162-15-8227 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 08 July 2020.

BOARD OF EXAMINERS



Dr. Syed Akhter Hossain
Professor and Head

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Chairman

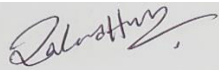


Dr. Sheak Rashed Haider Noori

Associate professor & Associate Head

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

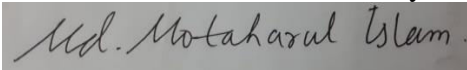


Md. Zahid Hasan

Assistant Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



Dr. Md. Motaharul Islam

Professor

Department of Computer Science and Engineering
United International University

External Examiner

DECLARATION

We hereby notify that, this project has been done by us under the supervision of **Raja Tariqul Hasan Tusher** 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:



Raja Tariqul Hasan Tusher
Sr. Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Aniruddha Rakshit
Sr. Lecturer
Department of CSE
Daffodil International University

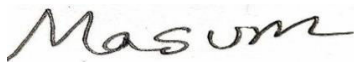
Submitted by:



Sadiqur Rahman
ID: 162-15-8181
Department of CSE
Daffodil International University



Alam Shahriar
ID: 162-15-8179
Department of CSE
Daffodil International University



Md. Alim-Ul-Alam Masum
ID: 162-15-8227
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

In the beginning, we grateful to almighty ALLAH for His divine blessing that makes us possible to complete the final year project successfully.

We are truly grateful and would like to wish our profound indebtedness to **Raja Tariqul Hasan Tusher** and **Aniruddha Rakshit**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “*Music-Source-Separation*” to implement this project. Their limitless motivation, professional counsel, constant surveillance, helpful guidance, positive mentality, beneficial assessment, continuous support have made this project possible.

From the depth of our heart, we would like to thank all of the faculty members of the Department of CSE for helping us in every possible way to finish our project. We would also thank every staff members of the CSE Department of Daffodil International University.

We are thanking our entire coursemate at Daffodil International University for taking part in this discussion during the completion of the course work.

Lastly, we thank our parents with gratitude for giving us their endless support and motivation.

ABSTRACT

In a world with full of sounds, music is used as a connection for people around the world. This proposed project titled "*Music-Source-Separation*" is a deep neural network reference implementation for creating opportunities for researchers, audio engineers and artists. The completion of the project will result in providing a genuine and clear concept of the sound and instruments. This project deposits music and permits users to separate pop music into four stems: vocals, drums, bass and the remaining other instruments. The user can observe the sound from the sequence that is inserted from the system. A lengthy history of music separation has a scientific interest because of being thought as an immensely difficult problem. For example, deep learning-based systems have been giving very meaningful separations which leads to increase interest commercially. Music-Source-Separation giving a reference implementation where deep neural network is basically established. The project itself provides two main purposes. Firstly, accelerating all academic research in this field. Secondly, Improving bengali music community. The Bengali music community has a bright history throughout the birth of the nation. Research for Bengali music has been very poor throughout the years. The artists also suffer from various problems during their careers and research has not been done for their work in their life. There is not a huge amount of collections from the past and also not enough resources created for future endeavors. Our work will try to affect the Bengali music community by creating a healthy and educative environment for all ages of people.

TABLE OF CONTENTS

CONTENTS	PAGE
Approval	ii
Declaration	iii
Acknowledgements	Iv
Abstract	V
CHAPTER 1: INTRODUCTION	1-2
1.1 Introduction	1
1.2 Motivation	1
1.3 Objectives	1
1.4 Expected Outcome	2
1.5 Report Layout	2
CHAPTER 2: BACKGROUND STUDY	3-4
	vi

2.1 Introduction	3
2.2 Literature Review	3-4
CHAPTER 3: METHODOLOGY	5-7
3.1 Introduction	5
3.2 Model	5
3.3 Input layer	6
3.3.1 Dataset	6
3.3.2 Dimensionality Reduction	7
3.3.3 Bi-directional LSTM	7
3.3.4 Output Stage	7
3.5 Separation	7
CHAPTER 4: IMPLEMENTATION AND TESTING	8-14
4.1 Installation & Imports	8-9
4.2 Uploading Dataset	9
4.3 The Target	10
4.4 The Spectograms	11-12
4.5 Separating The Signals	

4.6 Tracks After separation	13
4.7 Evaluating using museval	13
	14
CHAPTER 5: DESIGN SPECIFICATION AND RESULTS	15-16
5.1 Result Of All Metrics,Targets And Method	15
5.2 SDR vocal	16
CHAPTER 6: CONCLUSION AND FUTURE SCOPE	17-18
6.1 Discussion and Conclusion	17
6.2 Goal	17
6.3 Future Scope	17
6.4 Challenges	18
References	18-20

LIST OF FIGURES

FIGURES	PAGE
Figure 3.1 Model	5
Figure 4.1: Installation	8
Figure 4.2: Imports	9
Figure 4.3: Uploading musdb dataset	9
Figure 4.4: Target values	8
Figure 4.5: Importing librosa and showing the spectrograms	9
Figure 4.6: Vocal spectrograms	9
Figure 4.7: Accompaniment spectrograms	10
Figure 4.8: Separating the signals	10
Figure 4.9: Track Playing after separation	11
Figure 4.10: Evaluating using musval	11
Figure 5.1: Result of all metrics,targets and vocals	12
Figure 5.2: Vocal SDR	22

CHAPTER 1

INTRODUCTION

1.1 Introduction:

Music's sources separation is described as a problem that has been fascinating to the researchers for many years. Our project entitled " Music-Source-Separation " can be a solution to the Bengali music community. Music connects people of all ages everywhere in the world. Quality sound can provide a genuine thought about music and sound to improve the atmosphere of music in our country. It will also help to gain our reputation in music in the past and will change the course of music in the future for good.

1.2 Motivation:

Music has a great influence on people for many years. By music's sources separation, we can easily identify the singers vocal. During music composition, we can modify vocal and other accompaniments. We are highly motivated for establishing this project for a healthy environment in the music industry in which people can engage themselves in better quality music

1.3 Objectives:

- The primary objective of this project is the separation of instrumental sound such as bass, vocal, drums etc. from a song.
- Creating an initiative for the betterment of Bengali music.
- Giving a clear concept of instruments and components to users.
- Developing a suitable and sustainable system.
- Strengthen the music community of our country.

1.4 Expected Outcomes:

- Accelerating academic research in music.
- Improving bengali music community.
- It will keep secure original music sources.
- It will remove additional sounds from original sound
- Modifying music's sound.
- It can be used for any other audio sounds.
- It will remove unexpected noise from main sounds.

1.5 Report Layout:

Chapter 1: Introduction- This chapter includes introduction, motivation, objectives and expected outcome of the project work we have done is written here and the report layout.

Chapter 2: Background- The following chapter describes everything of the background situation of our work that has been done. We also provided the literature review of project, relative studies and tests of the system.

Chapter 3: Project methodology – This chapter has discussed about the method we applied for the construction of the system. This sector has the methods and steps, data collection procedure, some numerical analysis of the cited system.

Chapter 4: Experimental results and discussion: - This chapter is discussed all the experimental result that has been attained by the proposed system is discussed along with the presentation analysis and a summary of the result is covered. Beside that we submitted the system design in this chapter.

Chapter 5: Conclusions- The last chapter contains not only the conclusion part but also shares some suggestion that will help this topic's further study.

CHAPTER 2

BACKGROUND STUDY

2.1 Introduction:

In this chapter we will discuss on various research activity. A lot of work has been done regarding this topic.

2.2 Literature Review:

There has been no existence of closed-form solution where many sources being registered into a signal either mono or stereo. For finding out the real challenge, researchers have utilized their supplementary skill about the signals and their recording and mixing. There is a huge quantity of methods that centers around method relating "classical" signal processing.

A huge quantity of methods being crafted by hand were tuned to a minimum quantity of music recordings. The comprehensive neutral assessment of these methods, however, was rarely practicable as datasets were not accessible in free and was not existing in those time. As a matter of fact, ground truth separated stems becomes beneficially helpful because it gives a significant assessment. Nevertheless, music used for business purpose naturally leads to copyright protection. Besides, the separated stems are thought like important resources in the industry of music recording but not available all the time.

In addition , some artists who select licenses like Creative Commons, credit goes to them. It gives permission of providing the stems, datasets delivered in the past five years were freely available along with data-driven methods that empowered their growth. From that time, progression is shown in performance. Stöter et al., (2019). The progress has been guiding to large presentation upgrade for which many audio tasks for instance Automatic Speech Recognition (ASR). In Automatic Speech Recognition (ASR) data was found available at a maximum rate Datasets containing speeches exceeding 10000 hours (Amodei et al 2016) were given permission in 2016. In this case, the permission was given by speech recognition community. For facing the obstacles, the suggested systems are very good in comparison to further methods. The systems gone through supervision, methods based on classical signal processing definitely exceeded methods based on

machine learning. As a result, those methods were very helpful as a rapid and sometimes easy to recognize baseline.

A group of the reference implementations of source separation will be described respectively. On the other hand, a few systems found commercially, such as Audionamix XTRAX STEMS, IZOTOPE RX 7 or AudioSourceRE, are not only recognized tools which will be found as open-source software but also be comfortable for experimentation research purposes.

The inaugural software, openBLISSART was available for general users for source separation. In 2011, it was released (Felix Weninger¹, 2011). The inaugural software was encoded in C++. It also stands responsible for non-negative matrix factorization(NMF) because of some classes of processes. In the year of 2012, Flexible Audio Source Separation Toolbox(FASST) had been introduced in (Alexey Ozerov, 2012), and encoded by the help of MATLAB/C++. Despite NMF methods supporting this, it also comprises other methods based on different models. At the year of 2016, a library entitled “The Untwist Library” was suggested in (Gerard Roma, 2016) which encompasses various methods, having a range from methods based on classical signal processing to feed-forward neural networks. Python 2.7 has been encoded in the library. Sadly, the update has not been done since the year of 2017. Moreover, automated testing has not been done to a huge quantity of it's methods. A framework “Nussl” which is recently introduced in (Ethan Manilow, 2018). It consists of a huge amount of methods normally gives focus on classical signal processing methods rather than techniques based on machine-learning. This particular framework includes a built-in interface which is used for frequent evaluation of metrics and data sets. A decent level of abstraction and well modularity is presented in this library. In spite of being challenging, the beginners can find it easy for them to determine the changing of the machine learning processes of the techniques.

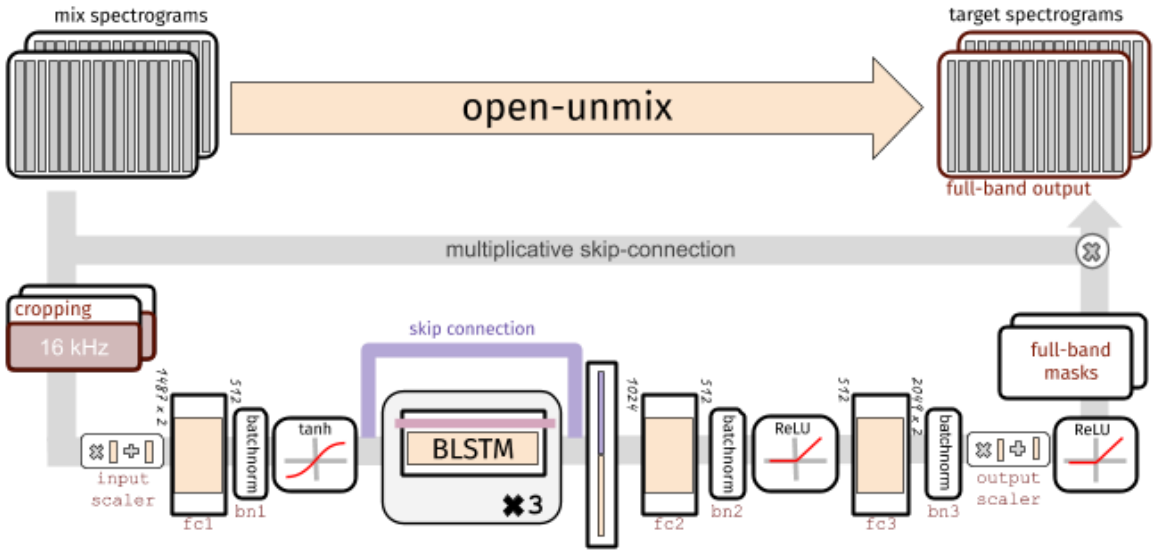
CHAPTER 3

METHODOLOGY

3.1 Introduction:

Music-Source-Separation is based on a three-layer bidirectional deep LSTM. This Model learns to predict the magnitude spectrogram of a target, like vocals, from the magnitude spectrogram of a mixture input. The prediction is obtained by applying a mask on the input. This model is optimized in the magnitude domain using mean squared error and the actual separation is done with a post-processing step involving a multichannel wiener filter implemented by norbert. Performing the separation into multiple sources, multiple models are trained for each particular target.

3.2 Model:



3.3 Input layer:

3.3.1 Dataset

The musdb18 is a dataset that embraces an amount of 150 full lengths music tracks of different styles. In this dataset, the music tracks' isolated drums, bass, vocals and other stems are included. The musdb18 carries two folders. One folder includes a training set composed of 100 songs. The other folder has a test set composed of 50 songs. The approaches should be trained on the training set which will be looked after. This will be tested on both sets. Every signals are stereophonic. They are encoded at 44.1 kHz. All files from the musdb18 dataset are encoded in the Native Instruments stems format (.mp4). This is a multitrack format composed of 5 stereo streams where each one is encoded in AAC @256kbps. The signals in those specific track leads to:

- 0 - The vocals,
- 1 - The drums,
- 2 - The bass,
- 3 - The rest of the accompaniment,
- 4 - The mixture.

We used Demo version of MUSDB18 which is 7s Samples .This demo version is included with musdb which is a python package to parse and process the MUSDB18 dataset.

Music-Source-Separation acts with the help of time-frequency domain in order to perform it's prediction.

A time domain signal tensor of shape (nb_samples, nb_channels, nb_timesteps), where nb_samples are the samples in a batch, nb_channels is 1 or 2 for mono or stereo audio, respectively, and nb_timesteps is the number of audio samples in the recording.

In that case, the model computes spectrograms with torch.STFT on the fly.

3.3.2 Dimensionality reduction

The LSTM is not operating on the original input spectrogram resolution. Instead, in the first step after the normalization, the network learns to compresses the frequency and channel axis of the model to reduce redundancy and make the model converge faster.

3.3.3 Bidirectional-LSTM

The core of Music-Source-Separation is a three layer bidirectional LSTM network. Because of having repetitive nature, the model's audio signals can be trained and appraised on arbitrary length. Though the model holds information from past and future in a simultaneous way, the model cannot be utilised in an online/real time manner. An uni-directional model can easily be trained as described here.

3.3.4 Output Stage

After applying the LSTM, the signal is decoded back to its original input dimensionality. In the last steps the output is multiplied with the input magnitude spectrogram, so that the models is asked to learn a mask.

$$Mask_{vocals} = \frac{S_{vocals}}{S_{vocals} + S_{acompaniment}}$$

3.3.5 Separation

Since PyTorch currently lacks an invertible STFT, the synthesis is performed in numpy. For inference, we rely on an implementation of a multichannel Wiener filter which is considered as a well admired way of filtering multichannel audio for some applications, particularly source separation along with speech enhancement. The norbert module assumes to have some way of estimating power-spectrograms for all the audio sources (non-negative) composing a mixture.

CHAPTER 4

IMPLEMENTATION AND TESTING

Introduction:

In previous chapter we discussed about how our system will work. We have also discussed some others work also. This chapter will have discussed about the system procedure what we are used to develop this system.

4.1 Installation of Requirements and Imports:

```
!pip install pandas==0.25.1
Requirement already satisfied: pandas==0.25.1 in /usr/local/lib/python3.6/dist-packages (0.25.1)
Requirement already satisfied: python-dateutil>=2.6.1 in /usr/local/lib/python3.6/dist-packages (from pandas==0.25.1) (2.8.1)
Requirement already satisfied: numpy>=1.13.3 in /usr/local/lib/python3.6/dist-packages (from pandas==0.25.1) (1.18.4)
Requirement already satisfied: pytz>=2017.2 in /usr/local/lib/python3.6/dist-packages (from pandas==0.25.1) (2018.9)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.6/dist-packages (from python-dateutil>=2.6.1->pandas==0.25.1) (1.12.0)

!apt install -y libsndfile1 ffmpeg
!pip install musdb
!pip install museval
!pip install -q scikit_posthocs
!pip install librosa
```

Fig 4.1:Installation

```

from IPython.display import Audio, display
import urllib.request
import numpy as np
import scipy.stats
import seaborn as sns
import pandas as pd
import scikit_posthocs as sp
import matplotlib.pyplot as plt
from matplotlib import gridspec
from matplotlib.transforms import BlendedGenericTransform
import scikit_posthocs as sp
from urllib.request import urlopen
import warnings
warnings.simplefilter(action='ignore', category=FutureWarning)

```

Fig 4.2:Imports

4.2 Uploading Dataset:

```

] import musdb
  mus = musdb.DB(download=True)

```

Downloading MUSDB 7s Sample Dataset to /root/MUSDB18/MUSDB18-7...
Done!

```

] track = mus[42]
  print(track.name)
  Audio(track.audio.T, rate=track.rate)

```

James May - On The Line

▶ 0:00 / 0:06 ————— 🔊 ⋮

Fig 4.3:Uploading musdb dataset

4.3 The target

```
] track.targets
↳ OrderedDict([('vocals', vocals),
               ('drums', drums),
               ('bass', bass),
               ('other', other),
               ('accompaniment', bass+drums+other),
               ('linear_mixture', vocals+bass+drums+other)])

] vocals = track.targets['vocals']
  Audio(vocals.audio.T, rate=track.rate)
↳
```




Fig 4.4: The target vocals, drums, bass, other

4.4 The Spectrograms:

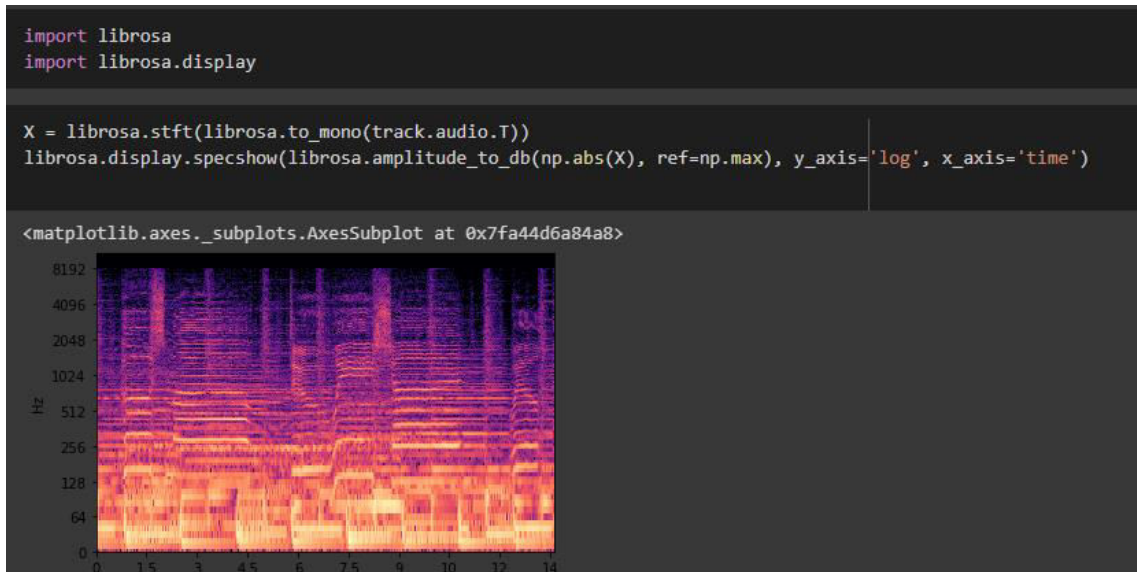


Fig 4.5: Importing librosa and showing the tracks spectrograms

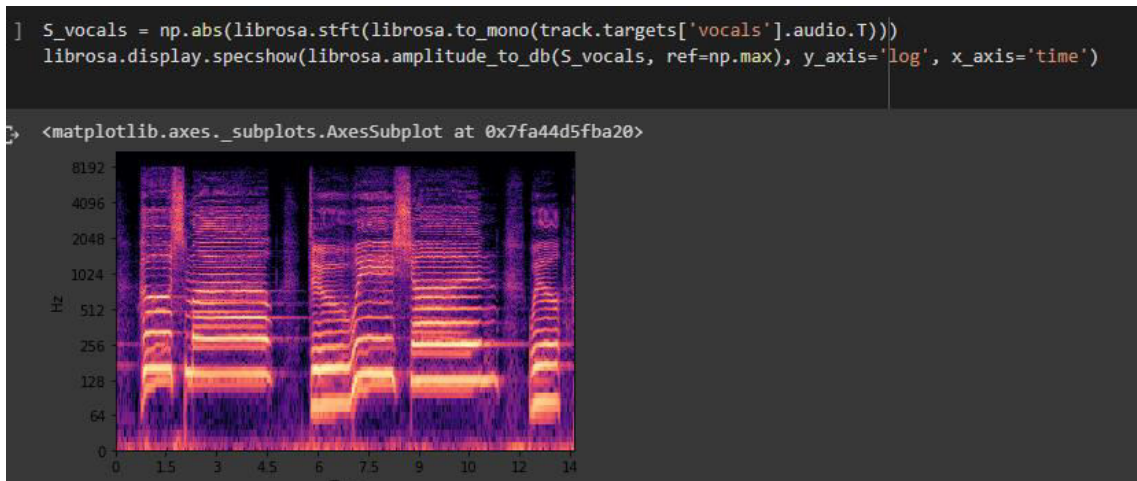


Fig 4.6: vocal spectrogram

```
S_accompaniment = np.abs(librosa.stft(librosa.to_mono(track.targets['accompaniment'].audio.T)))  
librosa.display.specshow(librosa.amplitude_to_db(S_accompaniment, ref=np.max), y_axis='log', x_axis='time')
```

<matplotlib.axes._subplots.AxesSubplot at 0x7fa44d5ff390>

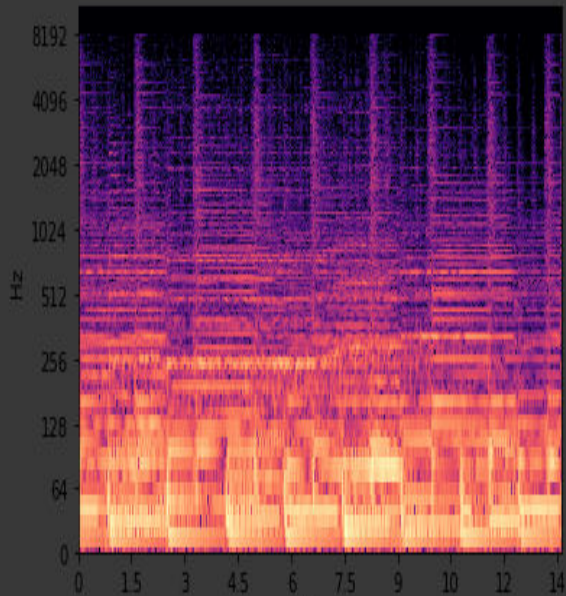


Fig 4.7:Accompeniment spectrogram

4.5 Separating The Signals:

```
import torch

eps = np.finfo(np.float).eps

N = track.audio.shape[0]
X = stft(track.audio.T, nperseg=4096, noverlap=3072)[-1]
(I, F, T) = X.shape

P = {}

model = eps

for name, source in track.sources.items():

    unmix = torch.hub.load('sigsep/open-unmix-pytorch', model='umx', target=name)
    P[name] = unmix(torch.tensor(track.audio.T[None, ...]).float()).detach().numpy()[:, 0, ...].transpose(1, 2, 0)

    model += P[name]

estimates = {}
for name, source in track.sources.items():

    Mask = P[name] / model

    Yj = Mask * X
    target_estimate = istft(Yj, nperseg=4096, noverlap=3072)[1].T

    estimates[name] = target_estimate

Downloading: "https://github.com/sigsep/open-unmix-pytorch/archive/master.zip" to /root/.cache/torch/hub/master.zip
Downloading: "https://zenodo.org/api/files/d6105b95-8c52-430c-84ce-bd14b803faaf/vocals-c8df74a5.pth" to /root/.cache/t
```

Fig 4.8: Separating the signals

4.6 Tracks After Separation:

```
[ ] for target, estimate in estimates.items():
    display(Audio(estimate.T, rate=track.rate))
```

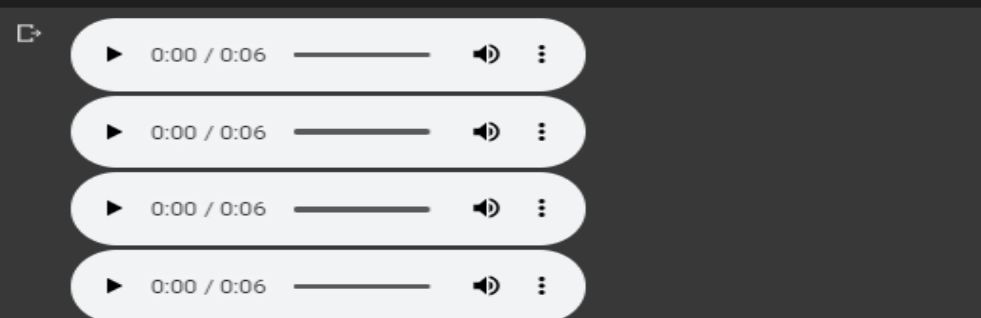


Fig 4.9: Track playing after separation

4.7 Evaluate Using museval

```
[ ] import museval

track_scores = museval.eval_mus_track(track, estimates)
print(track_scores)
```

```
vocals      ==> SDR: 8.310  SIR: 10.591  ISR: 13.464  SAR: 11.004
drums       ==> SDR: 6.793  SIR: 9.253   ISR: 9.555  SAR: 8.098
bass        ==> SDR: 5.393  SIR: 5.373   ISR: 9.896  SAR: 8.575
other       ==> SDR: 4.096  SIR: 7.660   ISR: 5.359  SAR: 5.740
```

track_scores.df

```
vocals      ==> SDR: 8.310  SIR: 10.591  ISR: 13.464  SAR: 11.004
drums       ==> SDR: 6.793  SIR: 9.253   ISR: 9.555  SAR: 8.098
bass        ==> SDR: 5.393  SIR: 5.373   ISR: 9.896  SAR: 8.575
other       ==> SDR: 4.096  SIR: 7.660   ISR: 5.359  SAR: 5.740
```

	time	target	metric	score	track
0	0.0	vocals	SDR	8.72872	James May - On The Line
1	1.0	vocals	SDR	10.07934	James May - On The Line
2	2.0	vocals	SDR	7.89189	James May - On The Line
3	3.0	vocals	SDR	6.94194	James May - On The Line
4	4.0	vocals	SDR	9.37876	James May - On The Line
...
91	1.0	other	SIR	5.78471	James May - On The Line
92	2.0	other	SIR	8.64014	James May - On The Line
93	3.0	other	SIR	6.67943	James May - On The Line
94	4.0	other	SIR	12.64383	James May - On The Line
95	5.0	other	SIR	8.88198	James May - On The Line

96 rows x 5 columns

Fig 4.10: Evaluating using museval

CHAPTER 5

5.1 Design Specification And Result

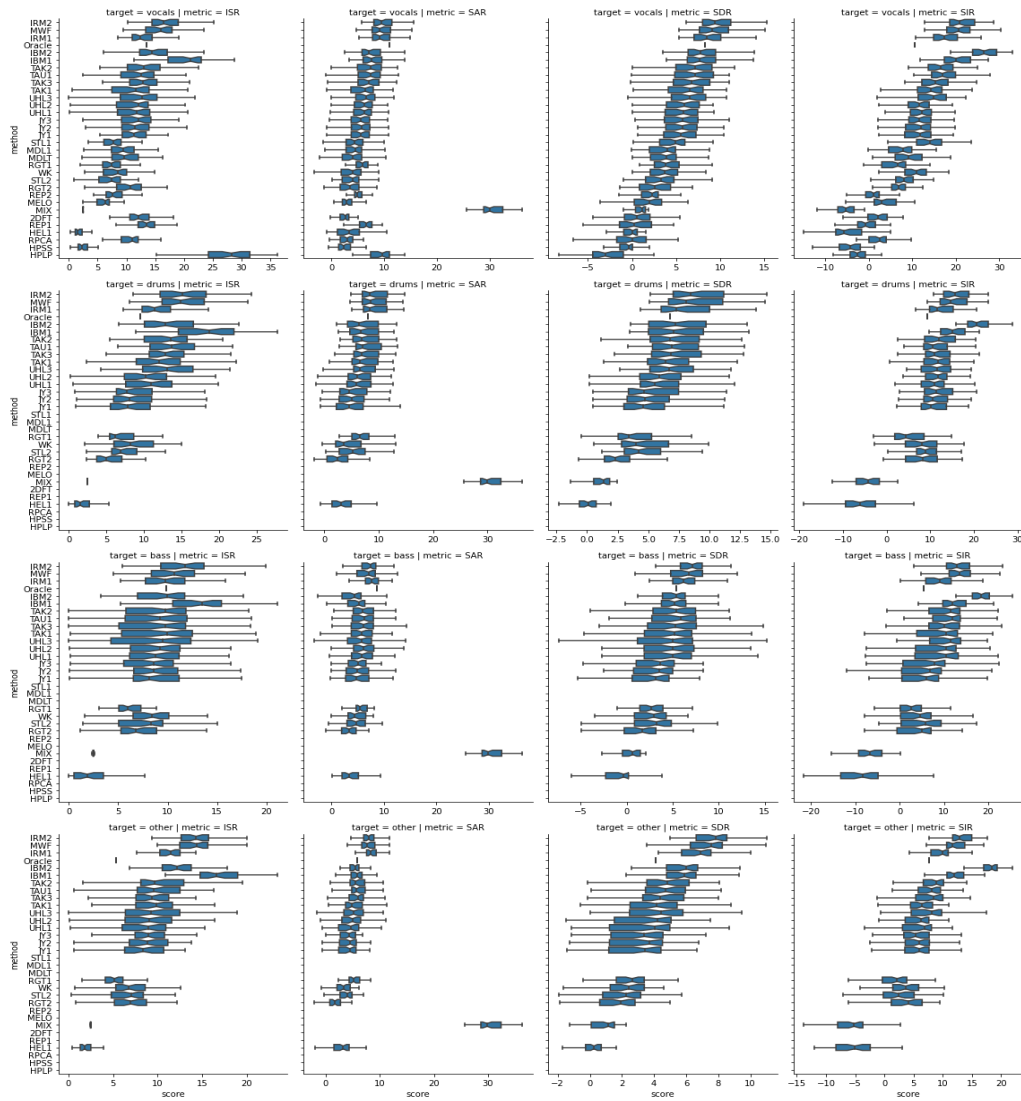


Fig 5.1: Details of results for all metrics, targets and methods.

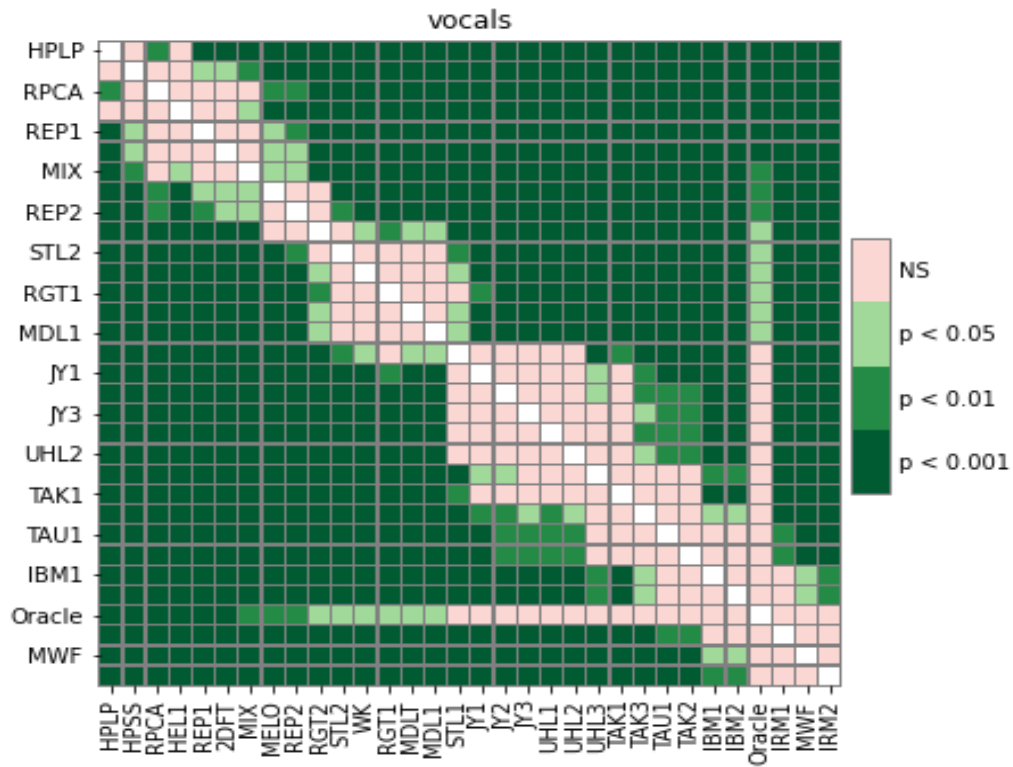


Fig 5.2: Vocal SDR

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

This project is ended up successfully with separating pop music into four stems: vocals, drums, bass and the remaining other instruments.

6.2 Goal:

- Developing a system that can easily separate pop music.

The system was designed to improve the Bengali music quality. It will also help gaining our reputation in music in the past and will change the course of music in the future for good .

6.3 Future Scopes:

We have developed a system which can work successfully. But we have decided in future, we will develop this system more efficient by following this step.

- If anyone try to improve Bengali music quality , it can be considered as a base.
- System will load and separate any kind of song or sound within a very short time .
- Loading full length songs easily.
- Making more user friendly.

6.4 Challenges:

We faced some problems with implementing our project.

1. Lacking of data.
2. Need Owner's permission for avoiding copyright problems.
3. Time consuming.
4. High speed internet access.
5. Need more GPU which is costly.
6. Need more time for training and testing model.

References

-] Ono, N. K. (2013).). The 2013 signal separation evaluation campaign. In Proc. *IEEE international workshop on machine learning for signal processing*(2013.6661988), 1–6).
- A. Amodei, D. A. (2016). Deep speech 2: End-to-end speech recognition in English and Mandarin. 173–182.
- Alexandre Défossez, N. U. (3 Sep 2019). Demucs: Deep Extractor for Music Sources with extra unlabeled data remixed. *1909.01174*.
- Alexey Ozerov, E. V. (2012). A General Flexible Framework for the Handling of Prior Information in Audio Source Separation. *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING*, 1118-1133.
- Araki, S. N. (2012). signal separation evaluation campaign (SiSEC2011): - audio source separation -. In 10th international conference on latent variable analysis and signal separation. .

- Bach, F. Y. (2014). Primal-Dual Algorithms for Non-negative Matrix Factorization with the Kullback-Leibler Divergence. *1412.1788*.
- Cano, E. F. (n.d.). Musical source separation: An introduction. *IEEE Signal Processing Magazine*, 36(1), 31–40.
- Chandna, P. M. (2017). Monoaural audio source separation using deep convolutional neural networks. In *Latent variable analysis and signal separation*. 258–266.
- Daniel Stoller, S. E. (6 Apr 2018). ADVERSARIAL SEMI-SUPERVISED AUDIO SOURCE SEPARATION APPLIED TO SINGING VOICE EXTRACTION. *1711.00048*.
- Daniel Stoller, S. E. (8 Jun 2018). WAVE-U-NET: A MULTI-SCALE NEURAL NETWORK FOR END-TO-END AUDIO SOURCE SEPARATION. *1806.03185*.
- David Samuel, A. G. (17 Feb 2020). META-LEARNING EXTRACTORS FOR MUSIC SOURCE SEPARATION. *2002.07016*.
- Ethan Manilow, P. S. (2018). THE NORTHWESTERN UNIVERSITY SOURCE SEPARATION LIBRARY. *Proceedings of the 19th ISMIR Conference, Paris, France*, 297-305.
- Fabian-Robert Stöter, S. U. (2019). Open-Unmix - A Reference Implementation for Music Source Separation. *The journal of open source software*, 01667.
- Felix Weninger¹, A. L. (2011). OPENBLISSART: DESIGN AND EVALUATION OF A RESEARCH TOOLKIT FOR. *Institute for Human-Machine Communication, Technische Universität München, Germany*, 1625-1628.
- Francesc Lluís, J. P. (28 Jun 2019). End-to-end music source separation: is it possible in the waveform domain? *1810.12187*.
- Gerard Roma, E. M. (2016). UNTWIST: A NEW TOOLBOX FOR AUDIO SOURCE SEPARATION. *Centre for Vision, Speech and Signal Processing, University of Surrey, UK*, 1-4.
- Joachim Muth, S. U. (16 Jul 2018). Improving DNN-based Music Source Separation using Phase Features. *arXiv:1807.02710v3*.
- McFee, B. K. (2018). Open-source practices for music signal processing research: Recommendations for transparent, sustainable, and reproducible audio research. *IEEE Signal Processing Magazine*, 36(1), 128–137.
- Michael Michelashvili, S. B. (6 May 2019). SEMI-SUPERVISED MONAURAL SINGING VOICE SEPARATION WITH A MASKING NETWORK TRAINED ON SYNTHETIC MIXTURES. *1812.06087*.

- Naoya Takahashi, Y. M. (29 Jun 2017). MULTI-SCALE MULTI-BAND DENSENETS FOR AUDIO SOURCE SEPARATION. *1706.09588*.
- Ozerov, A. V. (2011). A general flexible framework for the handling of prior information in audio source separation. . *IEEE Transactions on Audio, Speech, and Language Processing*, *20(4)*(10.1109/TASL.2011.2172425), 1118–1133.
- Romain Hennequin, A. K. (2019). SPLEETER: A FAST AND STATE-OF-THE ART MUSIC SOURCE SEPARATION TOOL WITH PRE-TRAINED MODELS. *Extended abstract for Late-Breaking/Demo*.
- Venkatesh S. Kadandale, J. F. (23 Jun 2020). Multi-channel U-Net for Music Source Separation. *arXiv:2003.10414v2* .
- Yi Luo, N. M. (15 May 2019). Conv-TasNet: Surpassing Ideal Time-Frequency Magnitude Masking for Speech Separation. *1809.07454*.

Turnitin Originality Report

Processed on: 07-Jul-2020 23:45 +08
 ID: 1354638263
 Word Count: 2514
 Submitted: 1

Music-Source-Separation By
 Sadiqur Raman 162-15-8181

Similarity Index

24%

Similarity by Source

Internet Sources:	22%
Publications:	13%
Student Papers:	16%

[exclude quoted](#) [exclude bibliography](#) [exclude small matches](#) mode:

quickview (classic) report

Change mode

[print](#) [refresh](#) [download](#)

3% match (Internet from 20-Sep-2019)

<https://hal.inria.fr/hal-02190845/document>

2% match (student papers from 07-Apr-2018)

[Submitted to Daffodil International University on 2018-04-07](#)

2% match (Internet from 03-Oct-2019)

https://hal.inria.fr/hal-02293689v1/html_references

2% match (Internet from 25-Mar-2017)

<http://rdm.uva.es>

1% match (Internet from 10-Oct-2019)

<https://hal.inria.fr/hal-02293689/document>

1% match ()

<https://hal.archives-ouvertes.fr/hal-02277338>

1% match (student papers from 18-May-2020)

[Submitted to University of Hertfordshire on 2020-05-18](#)

1% match (student papers from 24-Nov-2014)

[Submitted to Tveder's Education Group on 2014-11-24](#)

1% match (Internet from 10-Sep-2019)

<https://hal-lirmm.crsd.cnrs.fr/lirmm-01766781/file/articlav3.pdf>

1% match (Internet from 11-Dec-2019)

<http://archives.lemic.net>

1% match (Internet from 30-Dec-2019)