DATASET

QawwalRang: An audio dataset for genre recognition of Qawwali

Anonymous

March 31, 2022

Abstract

A new audio dataset containing 72 Qawwali songs of one-minute duration with associated metadata is presented. Majority of samples in the dataset exhibit strong similarity in fundamental musical properties making it suitable for cross-cultural music information retrieval studies. This is demonstrated by using the dataset in a Qawwali recognizer experiment. Qawwali recognition is an unsupervised genre classification algorithm based on source separation of two major components typical of Qawwali performances namely Tabla and Taali. The proposed algorithm shows a mean *Accuracy* across popular western music genres of almost 90% with a *Recall* of 76% on Qawwali songs. Public release of the QawwalRang dataset is accompanied by software that can be used to reconstruct and extend the dataset as well as reproducing the genre classification results presented in this study.

Keywords: qawwali, musical dataset, music genre recognition

1. Introduction

Qawwali has a long and historic tradition in Indian subcontinental music, by some accounts it goes back to at least 13th century. One of the earliest Indian classical music families "Qawwal Baccha Ghrana" (Bhat, 2009) is credited with refining this singing form as a distinctive genre. Qawwali has been used to recite devotional songs, usually sung by Qawwal groups on shrines of religious figures across South Asia. Even today they are popular in Indian subcontinent with performances given on happy as well as sad occasions. They are part and parcel of pop culture both in India and Pakistan with songs of this genre often included in film, festivals and commercial music.

Qawwali performances typically incorporate both classical and semi-classical musical elements. Along with some other genres such as Khayal, Thumri or Ghazal, Qawwali belongs to a "light" classical (Qureshi et al., 1986) sub-category, while having the maximum flexibility for improvisation. Two perceptually differentiating factors of Qawwali performance are both related to rhythm and beat patterns. Like the majority of Indian classical music, the percussion instrument Tabla[1] supplies rhythm, but in the Qawwali genre the rhythm patterns (Taals) are quite limited in scope. E.g. most Qawwali songs employ Keherwa[2] and Dadra[3] Taal patterns with eight and six beats respectively. A secondary rhythm pattern within Qawwalis is led by backing vocals clapping their hands at a periodic interval named as Taali (not to be confused with Taals). Both Tabla and Taali combine to give Qawwalis their characteristic rhythm structure.

Music information retrieval studies targeting dataset(s) from the Indian subcontinent have been reported in much fewer numbers as compared to their western counterparts. In fact there is a real lack of publicly available audio/feature datasets from South Asian region. Out of the reported literature, (Gulati et al., 2014) evaluated the fundamental pitch of vocals as well individual instruments on six different datasets focusing solely on classical Indian music. Few efforts to categorize Indian music into genres can be found in internet[4]. The open-source MusicBrainz[5] database includes songs tagged as Qawwali but there is no real distinction between generalized Sufi music and Qawwali in this project. None of these datasets qualify as a source of ground-truth for Qawwali genre recognition. This paper introduces a new publicly released dataset containing Qawwali audio snippets suited for cross-cultural music information retrieval research.

Genre recognition is a well established domain within music information retrieval research. Sturm (2014) gives a comprehensive overview of datasets, design strategies and performance measures used in reported genre classification results. According to them, genre recognition on publicly available dataset has

mostly dealt with western music. Recently, genre classification of eastern music datasets has also been reported. For example, MIREX 2020 (Mirex, 2020) included a task to estimate K-Pop genre detection. One of the earliest studies on music genre recognition is by Tzanetakis and Cook (2002). They introduced the GTZAN dataset which labels 100 songs of 10 western music genres each; even with its shortcomings this dataset is still being used in genre classification research. The objective of this paper is to introduce a new "Qawwali" genre that can be used alongside datasets like GTZAN. The dataset is constructed in a way so that it can be used in multicultural music research. To this end, "QawwalRang" dataset is supported with an unsupervised Qawwali genre recognition algorithm reporting results against the 10 popular western music genres included in the GTZAN dataset.

Rest of this paper is organized as follows: Next section describes the process of creating "QawwalRang" dataset, selection criteria to choose songs and a distribution of songs by artist and source of origin. Section 3 proposes an unsupervised classification system to recognize Qawwali songs. It details the features used to individually detect Tabla and Taali components before producing a binary classification label. Section 4 presents the results of applying genre classification on QawwalRang and GTZAN datasets. Observations on which western musical genre(s) may be more similar to Qawwali as compared to others are also included in this section. Section 5 covers the software used to build the dataset as well as an implementation of the classification algorithm. We end with concluding remarks and a list of planned future work.

2. QawwalRang Dataset

The QawwalRang dataset is a collection of 72 expert-labeled Qawwali songs. Each song, typically, is an excerpt from a much longer original recording. This excerpt is re-sampled at 44.1 KHz and converted into a single (mono) channel. The resulting song is tagged with a set of metadata values capturing its origin and creation process as detailed below.

2.1 Selection

Though Qawwali is an old and popular genre, the flexibility it offers, makes ground-truth creation of Qawwali songs a non-trivial task. This is partly due to the close association of Qawwali with so-called "sufi" music(Qureshi et al., 1986) which can be also be performed in other genres like Folk or Bhajan. Additionally, incorrect labeling in online databases is presumably due to the improvisation aspect; artists can mix/match Qawwali styles with modern music forms like pop music. In order to achieve a meaningful ground-truth level, QawwalRang construction inducted songs with the following criteria in mind:

- Song is performed by well-known/renowned Qawwal artists/groups (also called Qawwal par-

ties). While collecting items for this dataset, one of the aims was to stay as close as possible to the original Qawwali form, while achieving sufficient artist diversity. Including songs from artists associated primarily with the Qawwali genre reduces the sample space and makes the labeling obvious both structurally and perceptually.

- Songs must have an introductory section including both essential rhythm elements namely Tabla and Taali. These two components make Qawwali perceptually distinct as compared to other musical genres of the Indian subcontinent and can be easily identified by a listening expert. In a very small minority of cases where songs start with a vocal/instrument improvisation but does contain segments with Tabla and Taali components at a later time-point, an offset is incorporated to skip to the time-segments containing Tabla and Taali components.

- Songs performed in classical/semi-classical style have been preferred, this reduces overlap with other genres like pop, folk and film music. Songs that follow a modern mixture of Qawwali, overlapping traditional Tabla/Taali components with mainstream instruments are avoided.

After a song fulfilling the above criteria is selected, the next step is to create a metadata entry for it. The metadata entry follows a key-value interface which defines the following keys:

1. ID: unique identifier of the song in the data-set
2. Name: song's descriptive name suitable for searching in global databases
3. Artist: performer's or Qawwal party's name
4. Location: URL of the original song from which the excerpt has been taken. For songs from personal collection, samples are uploaded to an accessible place in the cloud and a URL provided. This is important since a Qawwali with the same name may bring up performances from many artists in an online search. Having a URL removes this ambiguity.
5. Start: Time-stamp in seconds within original song from where snippet was taken
6. Duration: Duration of song snippet in seconds, default is 1-minute. The default value is chosen with the consideration that some Qawwali performances can easily have a running time of 20-30 minutes.

As a result, construction of the QawwalRang dataset produced an associated metadata structure which completely specifies the dataset. Users can rebuild the dataset or even a modified version of it with software parsing this metadata structure. Please refer to the Reproducibility section for details on how this can be achieved.

2.2 Sources

Initial attempt to build the QawwalRang dataset focused solely on sourcing songs from personal collections. This was in part motivated by the first point in the selection criteria, namely to include as many classical Qawwali songs as possible. However, poor recording quality, non-availability of old recordings and the effort of ripping material off from cassettes/compact-disks and conversion to digital format proved quite daunting. To overcome these challenges we turned to the internet where a huge cache of songs is available, albeit, the ratio of songs fulfilling the selection criteria to the total number of Qawwali songs returned by a search engine is surprisingly low. Accepting songs from an online service has the advantage, though, of removing the ambiguity as to which particular performance of that song has been included in the dataset. Eventually QawwalRang songs originate from three sources i.e. internet music sharing platform(s), among others, YouTube[6], author(s) personal collection and crawling the web for Qawwali songs. Table 1 shows the distribution of sources in the QawwalRang dataset.

Source	Dataset Share(%)
YouTube	61
Personal	30
Web-Crawler	9

Table 1: QawwalRang: Source Distribution

2.3 Diversity

During compilation of the presented dataset, every attempt has been made to include songs from artists with different styles, active-periods and musical preferences. Figure 1 shows artist distribution in the QawwalRang dataset. This may seem skewed at first, but among the selected artists Nusrat Fateh Ali Khan (1948-1997) (MusicBrainz, 2022) and his group contribute a major share. Readers familiar with the Qawwali genre will probably know that Nusrat and his group were the foremost exponents of this genre in the last quarter of the previous century. They popularized the Qawwali genre with tireless performances across the world throughout the 70s/80s/90s and have remained widely popular and relevant even in recent times. In that sense their big share in the dataset is merited. The next big contribution comes from the 'others' category which is actually 9 songs each from a different artist. Rest of the dataset is quite evenly distributed among other selected artists which brings the total number of artists in the dataset to be 23.

3. Qawwali Genre Detector

The main motivation for developing a Qawwali recognizer is to evaluate suitability of the QawwalRang dataset for music information retrieval (MIR) research. We wanted to study fundamental signal properties of
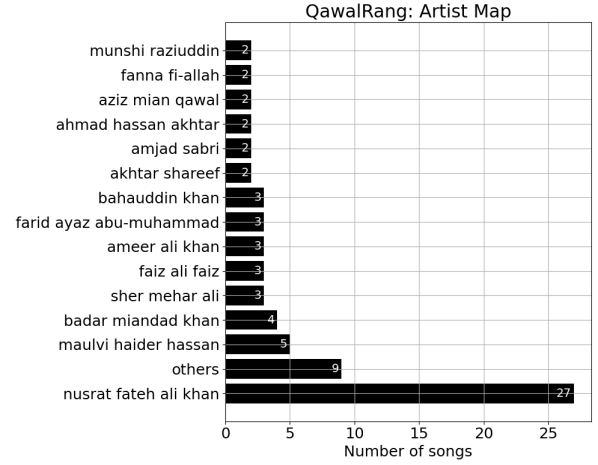

Figure 1: QawwalRang: Artist Distribution

the Qawwali samples; to be able to draw parallels between Qawwali and western music genres. Two of the most common signal properties used in music genre detection algorithms are Constant-Q Transform (CQT) and Mel-Frequency Cepstral Coefficients (MFCC) (Panagakis et al., 2014). Former is the time-frequency representation of an audio signal with frequency separation modeled on human ear, while the later captures the timbre quality of the signal. Instead of directly extracting these properties or features (as they are sometimes referred to) from polyphonic audio samples in the dataset, we factor in the characteristic Qawwali rhythm components namely Tabla and Taali. The idea is to first extract these components from polyphonic audio using non-negative matrix factorization, an algorithm suitable for unsupervised sound source separation (Virtanen, 2007). Tabla, being the main percussion instrument and Qawwali supporting limited Taals, is expected to contribute a big chunk of spectral energy in lower octaves. This hints towards the suitability of the CQT feature extracted from Tabla separated sound. Taali on the other hand is a short-time transient event without a fixed spectral distribution, while contributing a unique timbre to the Qawwali sound. Thus, MFCC seems like an appropriate feature choice for Taali separated source.

The proposed Qawwali genre classification thus consists of two stages, a feature extraction and a binary classification stage. The feature extractor separates Tabla and Taali components from the raw audio data then extracts CQT and MFCC from each component respectively. Individual decisions on Tabla and Taali detection are based on heuristics for CQT pitch energy and local MFCC extrema points. This is then fed to a binary classifier which decides if it recognizes the song as Qawwali. The overall scheme is shown in Figure 2
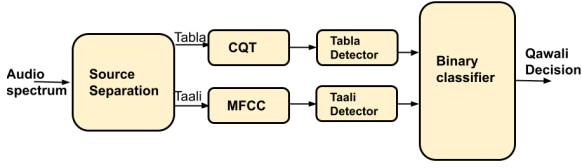
Figure 2: Proposed Qawwali recognition system.

3.1 Feature extractor

Features are extracted from raw audio spectrogram after separating the audio in potential Tabla/Taali source spectrograms. Let S be an $t \times m$ audio spectrum matrix decomposed into sorted feature matrix W and coefficients matrix H each of dimensions $t \times n$ and $n \times m$ respectively with non-negative matrix factorization, where n is the number of independent music sources the song is composed of.

$$S = W.H \quad (1)$$

We select $l \in (0, n)$ to extract the Tabla spectrogram S_B

$$S_B = W_B.H_B \quad (2)$$

where

$$W_B = [W.u_l ...] \quad (3)$$

and

$$H_B = [u_l^T.H ...] \quad (4)$$

with u_l being a unit column vector from the Identity matrix I with size $t \times t$. In other words the spectrogram for Tabla source is computed by selecting columns from the feature matrix H and rows from the coefficient matrix W. Lower column vectors are chosen for Tabla while choosing higher column vectors for Taali spectrogram S_T expressed as:

$$S_T = W_T.H_T \quad (5)$$

For Tabla detection, the extracted feature is CQT energy, obtained by mapping CQT from the Tabla spectrogram and taking norm/energy along the second dimension. This transforms the Tabla spectrogram matrix into a feature vector represented by:

$$f_{CQT} = ||f : S_B \rightarrow CQT|| \quad (6)$$

For Taali detection, a median MFCC vector is computed by obtaining a normalized MFCC vector from the Taali spectrogram as follows:

$$f_{MFCC} = med(\frac{f : S_T \rightarrow MFCC}{||f : S_T \rightarrow MFCC||}) \quad (7)$$

3.2 Binary Classifier

Qawwali recognizer is a binary classifier labeling each song either as Qawwali or non-Qawwali. It operates on the individual decisions produced by Tabla and Taali detectors, which in turn are based on extracted CQT and MFCC features. Below is a summary of individual component detection:

- Tabla detector (**TD**): Does f_{CQT} represent a Tabla source? Result is a non-binary decision with three possibilities *Yes(Y)/No(N)/Potential(P)*. The rationale for allowing an indeterminate decision at component level is that Tabla can be tuned to a different tonic, resulting in a shift of CQT energy.
- TaaLi detector (**TL**): Does f_{MFCC} represent a Taali source? This is based on extrema points within MFCC vector. Result is a binary *Yes(Y_l)/No(N_l)* decision.

Overall binary classification scheme is shown in the decision-tree diagram below, where leaf nodes **Q** indicate a decision in favor of Qawwali and **NQ** is the case when Qawwali genre was not recognized. Essentially it means the binary classifier categorizes a song as Qawwali only in two cases: first one where both Tabla and Taali were detected and an additional case where Taali has been detected from MFCC but Tabla detection from CQT was a "potential" call.

```
                    TD
          ┌─────────┼─────────┐
          Y         N         P
         / \       / \       / \
     TL:Y_l TL:N_l TL:Y_l TL:N_l TL:Y_l TL:N_l
       |     |     |     |     |     |
       Q    NQ    NQ    NQ     Q    NQ
```
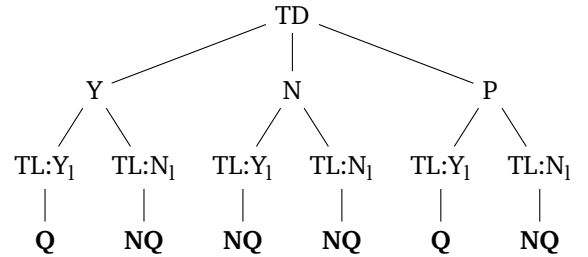
Tabla detection attempts to fit a Gaussian distribution onto energy in CQT bins as shown below:

$$f_{CQT} \sim \mathcal{N}(\mu, \sigma^2) \quad (8)$$

where $\mu \in [p_1, p_2]$ is the mean pitch within an octave band and $\sigma^2 < T_h$ is the variance of pitch (measured in musical notes). This is akin to finding a peak in the CQT energy profile, albeit giving weight to neighboring bins as well. p_1, p_2 are fixed parameters chosen at the start of the classification process whereas threshold T_h is a tunable parameter indicating bandwidth of notes containing a large concentration of CQT energy. Tabla is positively detected if a Gaussian curve successfully fits with $\mu \in [p_1, e_n]$ with a reasonably small threshold value. A potential decision for Tabla detection is made in case of $\mu \in [e_n, p_2]$, where a larger value of threshold is tolerated with e_n being edge-pitch parameter tunable during experiments. Failure to fit a distribution in either of these two ranges results in a negative Tabla decision.

For Taali detection, fitting a distribution is not plausible due to the much lower number of coefficients as

compared to CQT bins. Here, we attempt to find sudden change in the MFCC vector by considering short bursts of hand clapping. Given M element MFCC vector with MFCC-index $i \in (0, M-1)$, Taali detector searches a local extrema point i.e. either $f_{MFCC}(i) \leq f_{MFCC}(j)$ or $f_{MFCC}(i) \geq f_{MFCC}(j)$ where j is a tunable parameter. The Taali decision is positive in case local extrema is found at a certain j index.

4. Results & Discussion

Experimental results for the Qawwali classifier have been obtained with two sets of parameters: fixed and tunable. Among fixed parameters is the number of independent sources assumed to constitute a Qawwali, expressed by n in (1). It is assumed that each performance in the QawwalRang dataset contains 4 independent sources namely Vocals, Harmonium, Tabla and Taali, out of which the last two are of interest to the proposed Qawwali classifier. Tabla source in (2) is constructed with first and second column vectors of the feature matrix while second and fourth column vectors constitute Taali source in (5). Second fixed parameter of the experiments is the number of CQT bins in (6) which is fixed at 84 ranging between musical notes (midi-numbers) C1(24) and C8(108). Tabla detection is based on CQT measurements between third and fourth octaves which means p_1 and p_2 in (8) are set to C3(50) and G4(67) respectively. The last fixed parameter is the number of MFCC used in Taali detection meaning f_{MFCC} in (7) is a one dimensional vector with $M{=}13$ coefficients. Source separation, CQT and MFCC are computed with corresponding functions from librosa (McFee et al., 2022), while curve fitting on Tabla CQT to find a peak of CQT energy in third and fourth octave has been done with the help of lmfit optimization library (Newville et al., 2014).

The tunable parameters of the experiments (with reference to previous section) are: edge-note within third octave (e_n), CQT energy threshold in third (T_{3h}) and fourth octave (T_{4h}) for Tabla detection and index (j) of local MFCC extrema for Taali detection. The proposed Qawwali classifier is governed by these four design parameters. In this section we discuss the impact of each of these parameters in detail leading to a set of values used for classifying Qawwali against other genres in the GTZAN dataset. An assessment on each of these parameters in studied in light of four figures of merit (FOM) (Sturm, 2014): *Accuracy, Recall, F-measure* and *Precision*.

The main objective of these experiments is to find suitable values for tunable parameters such that proposed Qawwali classifier is able to separate songs in QawwalRang dataset from the western music samples in GTZAN dataset; properly tuned parameters should also produce results offering insights into similarities between western music genre(s) and Qawwali. Figure 3 shows the FOM plotted against various values

of edge-notes within the third octave. In order to drive these results, other variable parameters are given some default values, energy threshold in third and fourth octaves is set to 3 and 1 respectively while local extrema point for Taali detection is assumed to be the middle point. We see that all four metrics saturate beyond midi-number 53 (F3) which effectively indicates an interval of [C3,F3] within the third octave. Recalling the binary classifier operation from the previous section, this means that the Qawwali recognizer should attempt to find a peak in CQT energy within this interval for positively detecting Tabla.

Figure 4 shows FOM plot against a threshold of CQT energy spread within the third octave. Here, we observe that a threshold of energy spread higher than 4 does not improve *Recall* (positive Qawwali detection) but reduces *Precision* and *Accuracy*, this is obvious because choosing a larger energy spread as threshold increased False Positives by pushing more songs from other genres to be detected as Qawwalis due to wrongly detecting Tabla presence. Last two plots sug-
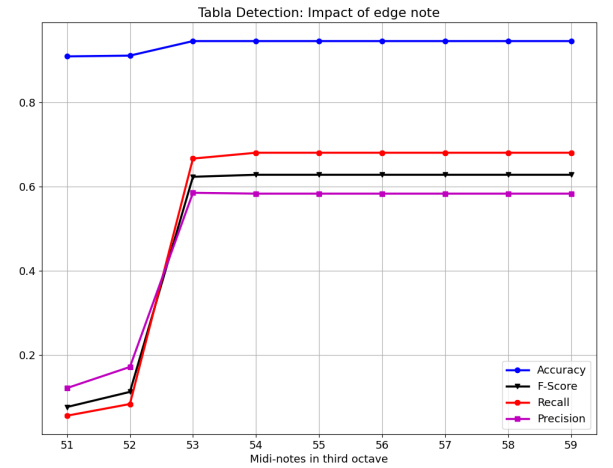

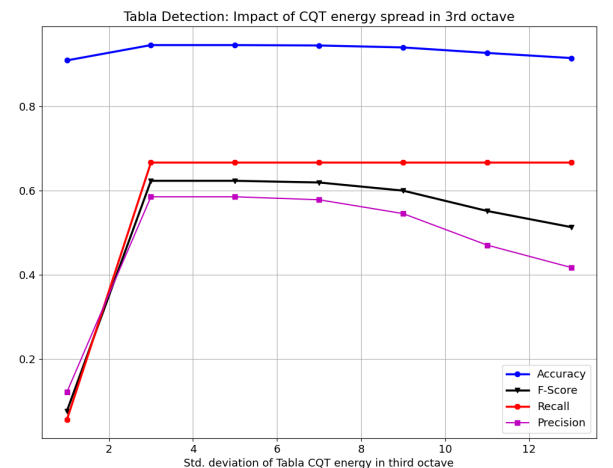**Figure 3:** Tabla Detection: Edge note in third octave

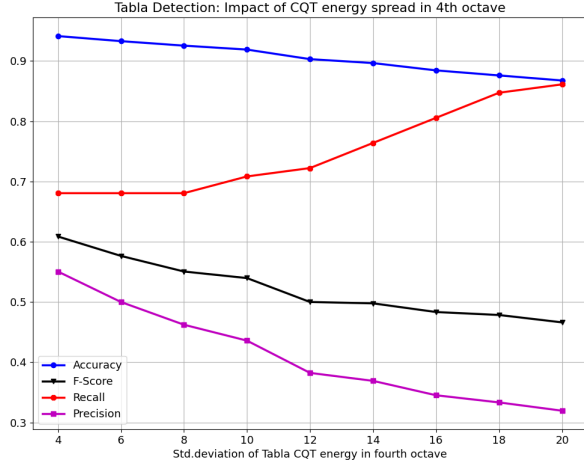**Figure 4:** Tabla Detection: CQT variance in third octave

Figure 5: Tabla Detection: CQT variance in fourth octave

Figure 6: Taali Detection: MFCC local extremum impact

gest that a good bet for the Qawwali detector is to decide in the favor of Tabla if it is able to fit a Gaussian distribution between [C3, F3] with variance under 4. However, as explained in the previous section, Tabla CQT energy concentration can vary anywhere between third and fourth octaves; a "Potential" decision is made in case an energy peak is detected within the fourth octave. Figure 5 shows the impact of choosing a CQT energy threshold for Tabla detection in fourth octave. We observe that selecting a higher threshold results in very good recall as almost all Qawwalis in the dataset exhibit Tabla CQT energy within the third or fourth octave. At the same time, the classifier identifies a lot of False positives with a large threshold. These experiments suggest a suitable value of 14 achieving a neutral F1-score.

The last tunable parameter is a local extrema within the MFCC vector for Taali detection. That is to say that the Qawwali detector decides in the favor of Taali if the local maximum/minimum value falls on a specific coefficient index (j). Figure 6 shows the impact on FOM upon choosing local extremum point to be one of 5th, 6th and 7th coefficient. Choosing another point resulted in vastly inferior performance results. Looking at the FOM metrics from Figure 6, 6th coefficient was chosen to identify the Taali source within the song.

Based on these results, tunable parameters for the Qawwali recognizer were determined to be: $e_n = 3$, $T_{3h} = 4$, $T_{4h} = 14$ and $j = 6$. Figure 7 shows the performance of the proposed Qawwali classifier using these parameters against music genres of GTZAN dataset. First thing to note is that a great majority of songs are correctly **NOT** classified as Qawwali. The mean *Accuracy* rate across GTZAN genres comes out to be 90.6% evaluated across 10 genres. At the same time a reasonable number of songs in the QawwalRang dataset are correctly identified as Qawwali with the *Recall* standing at 76%. Really interesting are the results for 'blues' and 'classical' genre(s) showing a high false positive
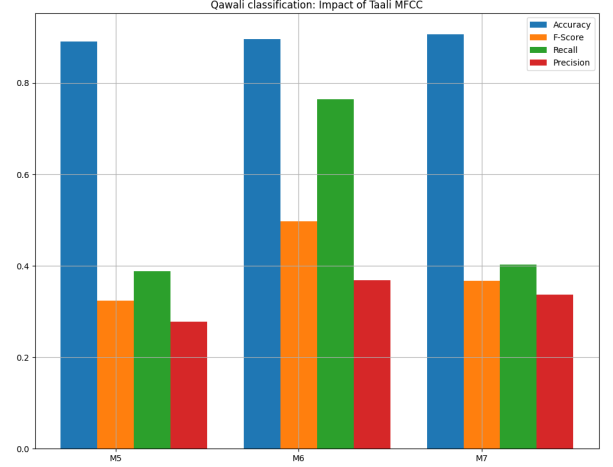
rate of 24% and 29% respectively. This suggests that in terms of fundamental music properties of CQT and MFCC, songs in these genres have a high likelihood to be identified as Qawwali. A closer look at the results from these genres indicate that the increase in False Positive cases is due to songs in these categories returning a positive Taali detection with a suspicion of Tabla. This indicates that at least in timbre quality songs from these genres would be perceived much closer to Qawwali as compared to other western genres. For the 'classical' genre this can be explained by the fact that the GTZAN dataset has many songs as acoustic piano instrumentals. This has the potential of overlapping with the harmonium, the leading melody instrument in Qawwali performances. For the 'blues' genre, characteristic blues rhythm sounding similar to the Taali structure in Qawwali is a possible explanation.
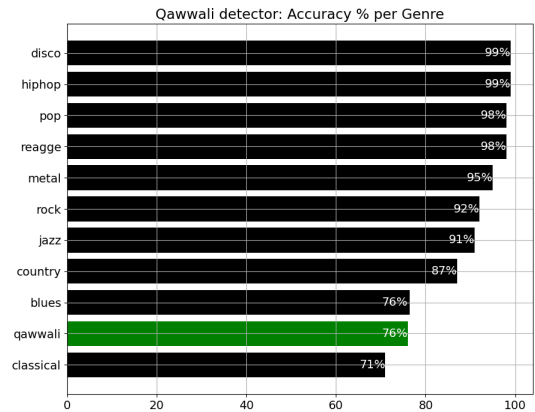

Figure 7: Qawwali detection per Genre

Perhaps a final noteworthy point about these results; they are not about a generalized genre detector being able to recognize a new Qawwali genre. In-

stead they highlight that QawwalRang songs possess fundamental musical properties making them easily recognizable against songs from western music genres. In that sense the proposed dataset can be employed by state of the art genre detection studies aiming to incorporate songs from non-western musical sources. What these results also imply is that with more intelligent feature extraction and/or classification techniques, there is a huge scope of obtaining detailed insights into cross-cultural music information retrieval, as long as appropriately labeled datasets are used.

5. Reproducibility

Accompanying the dataset are command-line interface (CLI) programs that can used to reproduce the results presented in this study. The software includes a python program to construct, extend or filter songs in the dataset. It operates on a JSON (Pezoa et al., 2016) formatted metadata file, editable by users before building a dataset. As explained in Section 2, the metadata interface allows the user to specify an offline songs directory/folder. This feature can be used to rapidly build a dataset from locally available songs. Song entries in metadata structure specifying online mode are downloaded, processed and then included in the resulting dataset. A python implementation of the proposed Qawwali recognizer is also part of the released software. It operates on the dataset constructed by the counterpart dataset builder program. The Qawwali recognizer program produces Qawwali labels per song with a summary at the end showing how many songs within the dataset were classified as Qawwali. This program can optionally be run in a feature-extraction mode, during which features are extracted and saved locally, skipping classification. This can be useful later with the program running the classifier on multiple feature-sets, each of which may belong to a different genre. The Qawwali detector program also supports a comparison option in which Qawwali classification results from various genres can be compared against each other. TODO: proper citation required DOI: https://github.com/fsheikh/QawwalRang

6. Conclusion & Future Work

An audio dataset of Qawwali songs along with a system to recognize the Qawwali genre has been presented. The developed system shows encouraging results when used in a cross-cultural genre classification study. Future work planned is: extending the diversity of dataset, genre recognition of Indian/Pakistani music genres other than Qawwali, and evaluation of QawwalRang with supervised machine learning techniques.

Notes

[1] Tabla
https://sriveenavani.com/Tabla
[2] Keherwa
https://www.taalgyan.com/taals/keherwa
[3] Dadra
https://www.taalgyan.com/taals/dadra/
[4] Indian Music Genre Dataset
https://www.kaggle.com/datasets/winchester19/indian-music-genre-dataset
[5] MusicBrainz
https://musicbrainz.org/tag/qawwali
[6] YouTube:
www.youtube.com

References

Bhat, J. T. (2009). *Hindustani Vocal Music*. Abhinav Publications New Delhi.

Gulati, S., Bellur, A., Salamon, J., H.G., R., Ishwar, V., Murthy, H. A., and Serra, X. (2014). Automatic tonic identification in indian art music: Approaches and evaluation. *Journal of New Music Research*, 43(1):53–71.

McFee, B., Metsai, A., McVicar, M., Balke, S., Thomé, C., Raffel, C., Zalkow, F., Malek, A., Dana, Lee, K., Nieto, O., Ellis, D., Mason, J., Battenberg, E., Seyfarth, S., Yamamoto, R., viktorandreevichmorozov, Choi, K., Moore, J., Bittner, R., Hidaka, S., Wei, Z., nullmightybofo, Weiss, A., Hereñú, D., Stöter, F.-R., Friesch, P., Vollrath, M., Kim, T., and Thassilo (2022). librosa/librosa: 0.9.1.

Mirex (2020). Audio k-pop genre classification. https://www.music-ir.org/mirex/wiki/2020:Audio_K-POP_Genre_Classification. Accessed: 2022-03-19.

MusicBrainz (2022). Nusrat fateh ali khan- musicbrainz. https://musicbrainz.org/artist/5968383c-11c1-4c5b-ada0-504a38cec8e7/. Accessed: 2022-03-19.

Newville, M., Stensitzki, T., Allen, D. B., and Ingargiola, A. (2014). LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python.

Panagakis, Y., Kotropoulos, C. L., and Arce, G. R. (2014). Music genre classification via joint sparse low-rank representation of audio features. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(12):1905–1917.

Pezoa, F., Reutter, J. L., Suarez, F., Ugarte, M., and Vrgoč, D. (2016). Foundations of json schema. In *Proceedings of the 25th International Conference on World Wide Web*, pages 263–273. International World Wide Web Conferences Steering Committee.

Qureshi, R., Qureshi, R. B., and Blacking, J. (1986). *Sufi music of India and Pakistan: sound, context and meaning in Qawwali*, volume 1. CUP Archive.

Sturm, B. L. (2014). A survey of evaluation in music genre recognition. *Adaptive Mutlimedia Retrieval: Semantics, Context, and Adaptation. Lecture Notes in Computer Science*, 8382:29–66.

Tzanetakis, G. and Cook, P. (2002). Musical genre clas-
    sification of audio signals. *IEEE Transactions on
    Speech and Audio Processing*, 10(5):293–302.
Virtanen, T. (2007). Monaural sound source separation
    by nonnegative matrix factorization with temporal
    continuity and sparseness criteria. *IEEE Transac-
    tions on Audio, Speech, and Language Processing*,
    15(3):1066–1074.