Analysis and Classification of Symbolic Western Classical Music by Composer

Stanford University

Saahil Jain, Akshay Smit, Tim Yngesjö {sj2675, akshaysm, timyn286}@stanford.edu

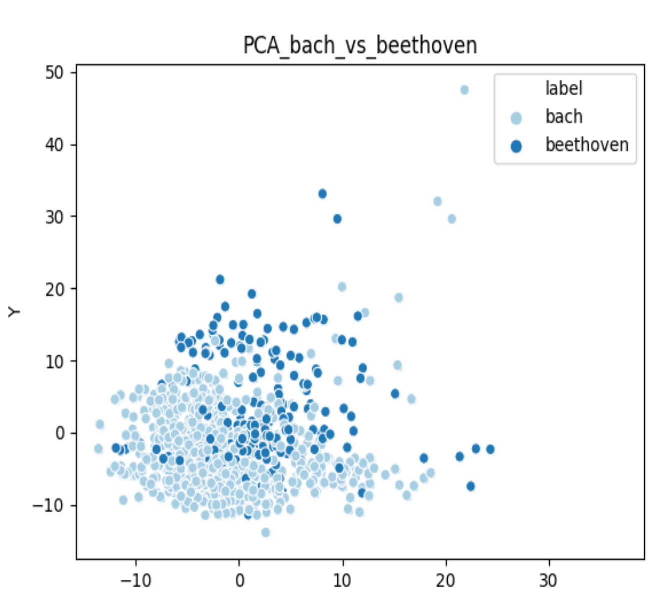
MOTIVATION

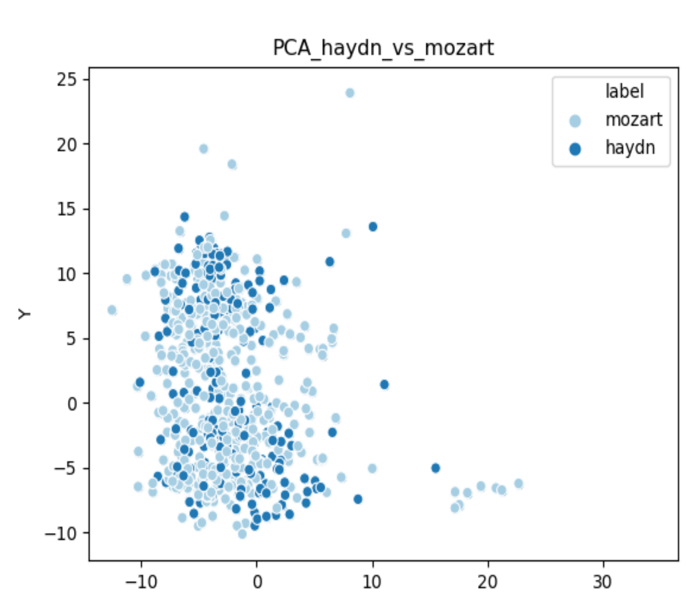
Our aim was to explore various methods of classifying pieces of Western classical music by composer and better understand the most distinctive features of such composers. Given the MIDI encoding of a musical piece, we convert the piece to a numerical feature vector and classify the composer. On 15 different composers, decision tree boosting achieved the highest test accuracy (79%) among all models. We also explored the potential of composer classification via computer vision, converting the MIDI encoding to an image and using CNNs to achieve 70% test accuracy on 6 composers. Additionally, we performed feature analysis, finding that vertical intervals and texture prove considerably more useful for classification than rhythmic features.

DATASET & FEATURES

We are very grateful to the Classical Archives, who generously provided us with over 10,000 MIDI files (representations of musical scores) for our project. We focused on 15 composers spanning the baroque, classical, and romantic eras of classical music (using 5,291 MIDI files). Using jSymbolic, an open source tool, we converted these MIDI files into 1,494-dimensional feature vectors, which numerically represent the rhythm, melody, texture, and tempo of a piece. For our image classification experiment, we convert each MIDI-file into a 64x1536 gray-scale image, where vertical axis corresponds to MIDI note, the horizontal axis corresponds to time, and the pixel value represents the velocity of the note if it is on. A small segment of a score generated from a MIDI file can be seen below:

Using PCA with two components, we can visualize musical pieces corresponding to composers and notice that composers from distinct eras (e.g. Bach & Beethoven on bottom left) appear more linearly separable than composers from similar eras (e.g. Haydn & Mozart on bottom right).





MODELS

We experimented with various models to classify the composers. The ones we considered were: **Gradient Boosting** (using decision stumps), **Fully-Connected Neural Networks, SVMs** with PCA (using a one vs. one scheme), **Softmax, K-Nearest Neighbors,** and **Linear Discriminant Analysis.**

Our top performing model, gradient boosting with decision stumps, uses the following initial state and update rule.

$$F_0(x) = \arg\min_{\gamma} \sum_{i=1}^{n} L(y_i, \gamma)$$

$$F_m(x) = F_{m-1}(x) + \arg\min_{h_m \in \mathcal{H}} \left[\sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + h_m(x_i))\right]$$

Where $h_m \in \mathcal{H}$ is the base decision stump and L is the loss function (in our case, L is the cross entropy loss).

CLASSIFICATION OF 15 COMPOSERS

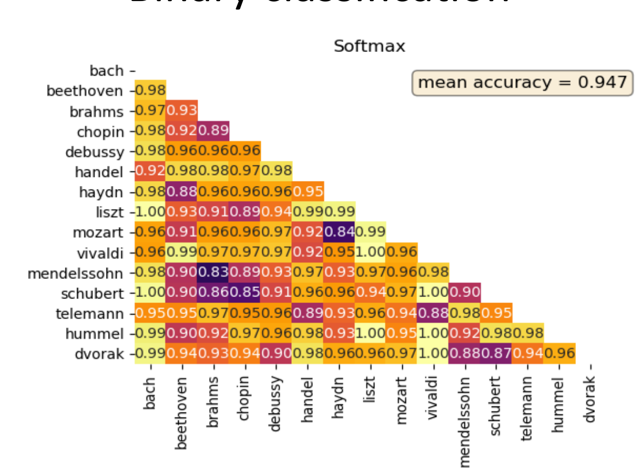
We provide accuracies and other metrics of other fine-tuned classifiers in the table below ranked from highest to lowest test accuracy. Gradient Boosting with decision stumps gave the highest accuracy (79.1%) on the test set.

Model	Train Accuracy	Dev Accuracy	Test Accuracy	Average Precision	Average Recall	Average F1 Score
GB	1.000	0.794	0.791	0.830	0.706	0.742
NN	0.989	0.752	0.752	0.720	0.684	0.698
Softmax	0.990	0.742	0.743	0.720	0.658	0.682
SVM	0.999	0.741	0.734	0.765	0.641	0.686
LDA	0.885	0.704	0.707	0.675	0.623	0.639
k-NN	0.648	0.564	0.575	0.617	0.433	0.463

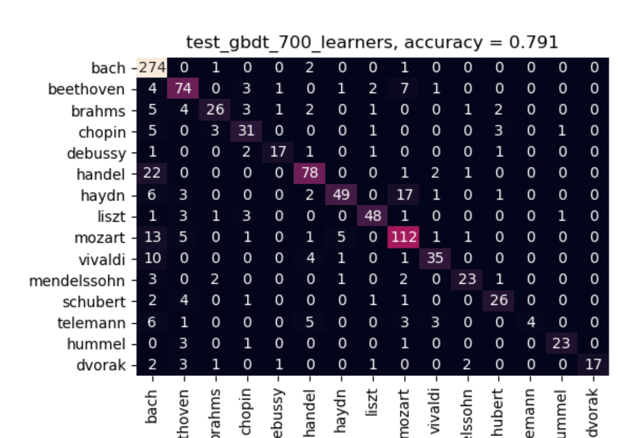
The confusion matrix for Gradient Boosting can be seen below in the bottom right figure. Composers from the same era are often confused, such as Handel and Bach (baroque) and Mozart and Haydn (classical).

Next, using softmax, we perform binary classification across each composer, which again shows that similar composers are often confused (bottom left figure).

Binary classification



Confusion matrix



Vertical intervals provide highest overall accuracy when classifying by era (as seen on the right). A mere 24 texture-related features provide more information than over 200 rhythmic features [using gradient boosting].

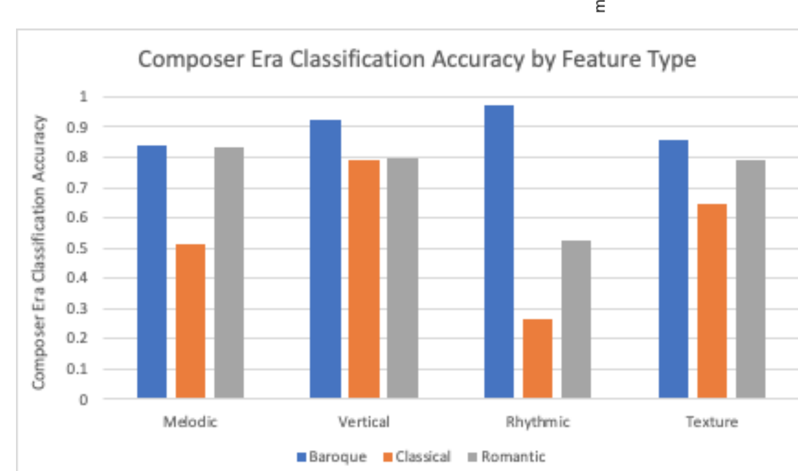
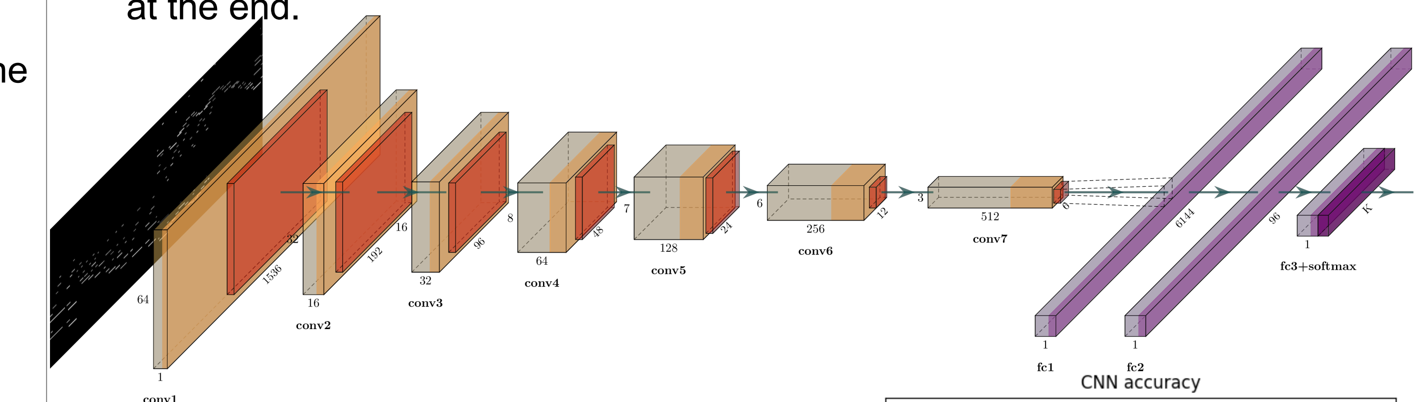
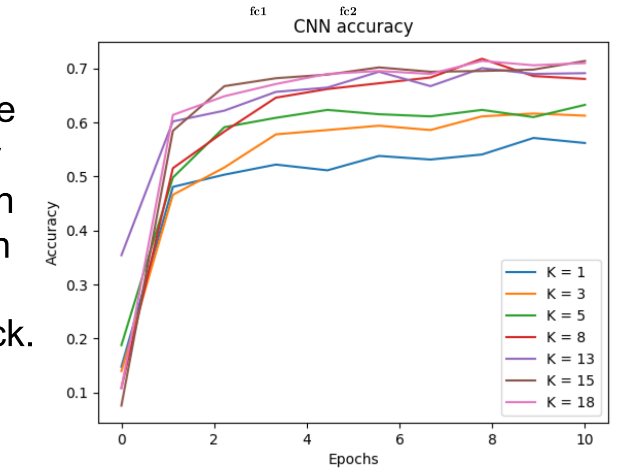


IMAGE CLASSIFICATION

Each MIDI-file is converted into a 64x1536 gray-scale image. A convolutional neural network (CNN) is then trained with the images as input. The CNN consists of 7 convolutional layers with max-pooling and 3 fully connected layers at the end.



On 6 composers the CNN achieved an accuracy of 70%. The performance of the CNN was improved from 55% to 70% by using data augmentation as can be seen in the plot to the right. The augmentation consisted of random translations and setting some blocks of the image to black.



DISTINCTIVE FEATURES

- The binary classification SVM was very good at separating composers from distinct eras. The top coefficients provide a measure of feature importances.
- Melodic and vertical intervals occur most frequently among the top coefficients for the binary SVM as well as gradient boosting, suggesting melody is more distinctive than rhythm.
- Dissonant melodic and vertical intervals (like tritones, dominant sevenths, second intervals) and large melodic range are the most distinctive features for romantic composers.
- Instrumentation is a very accurate giveaway. For instance, harpsichord immediately implies baroque, whereas a great deal of piano prevalence suggests a romantic composer (like Debussy or Chopin). French Horn is usually very indicative of Mozart.
- Use of standard triads and dominant spread (both melodic features) was very effective at differentiating baroque and romantic composers.

FUTURE WORK

- Further explore image classification by extracting more information from the MIDI-file to create a multichannel image instead of a grayscale image. Also investigate CRNNs.
- Generate music scores similar to a given composer using GANs and CRNNs.
- Generate sequences of notes using GANs and RNNs.

REFERENCES

Choi, K., Fazekas, G., Sandler, M., and Cho, K. Convolutional recurrent neural networks for music classification. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) , pp. 2392–2396, March 2017. doi: 10.1109/ICASSP.2017.7952585.

Yust, J., Wild, J., and Burgoyne, J. Mathematics and Computation in Music: 4th International Conference, MCM 2013, Montreal, Canada, June 12-14, 2013, Proceedings. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013.ISBN9783642393570

Microsoft. "Welcome to LightGBM's Documentation." Welcome to LightGBM's Documentation! - LightGBM 2.3.2 Documentation, lightgbm.readthedocs.io/en/latest/.