

Estimating Piano Piece Difficulty via Embodied Robotic Hand Performance Analysis

Alexander Neuhauser
Institute of Computational Perception
Johannes Kepler University Linz
Linz, Austria
alexander.neuhauser@jku.at

Carlos Cancino-Chacón
Institute of Computational Perception
Johannes Kepler University Linz
Linz, Austria
carlos.cancino_chacon@jku.at

Abstract

Piano performance combines precise motor coordination with artistic expression, making it a rich domain for studying complex, goal-directed movement. In this work, we explore how reinforcement learning can be used to model the embodied process of acquiring such piano-playing skills. Building on the ROBOPIANIST framework, a simulated system for piano performance, we focus on estimating the difficulty of musical pieces through the agent's learning dynamics rather than analyzing the structural properties of the musical score. Preliminary experiments show a statistically significant inverse correlation between the agent's learning performance and human-rated difficulty, indicating that embodied learning can serve as a proxy for estimating performance difficulty.

CCS Concepts

• **Computing methodologies** → **Reinforcement learning**; *Physical simulation*; • **Applied computing** → **Performing arts**; Sound and music computing; • **Computer systems organization** → *Robotic control*.

Keywords

piece difficulty estimation, music performance, reinforcement learning

ACM Reference Format:

Alexander Neuhauser and Carlos Cancino-Chacón. 2026. Estimating Piano Piece Difficulty via Embodied Robotic Hand Performance Analysis. In *MOCO '26: Proceedings of the 10th International Conference on Movement and Computing (MOCO '26), April 23–25, 2026, Montpellier, France*. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3802842.3802850>

1 Introduction

Since its invention more than three centuries ago, the piano has played a central role in Western musical culture. The structure of the keys enables a wide range of opportunities and the possibility to play multiple voices at the same time. This is one of the key reasons why the piano has become one of the most popular instruments and why many musicians, even those specializing in other instruments, learn to play it as part of their musical education [7]. Achieving expert proficiency typically requires years of deliberate practice, as

reflected in conservatory curricula and graded examination systems such as the ABRSM [1, 8, 35].

To play the piano successfully, a pianist has to deal with several challenges:

- (1) **Motor coordination:** A piano player has to be able to play with both hands simultaneously and independently, whereas each hand is responsible for playing a different set of keys. At the same time, the individual fingers should be able to move independently [9, 11].
- (2) **Rhythmic accuracy:** Precise and correct timing is another challenge the pianist has to achieve, which can be especially demanding when playing particular virtuosic musical passages [10, 11].
- (3) **Read sheet music:** In notated music genres like Western classical music, a pianist should be capable of reading (i.e., decoding) written musical scores. This process implies not only identifying the notes to be played in the musical score, but also mapping these notes to the corresponding keys in the keyboard and developing a motor program that translates these written-note-to-keys mappings into movement sequences [19, 24].
- (4) **Expression, Dynamics, and Musicality:** Merely playing the notes that have been written on the musical score in a mechanical way is not enough to convey a satisfactory musical interpretation. The pianist has to be able to understand and convey musical nuances in the musical structure by using patterns of variability in aspects such as dynamics, timing, and articulation [4, 17, 24]. The composer's intent should be interpreted, and the performer should be able to demonstrate a sense of musicality, which often requires a deep musical understanding and years of practice.

Although learning any piano piece poses its own challenges, the level of complexity can vary a lot from piece to piece, and recommending students pieces to learn that match their skill level requires a lot of implicit expertise from music teachers. Judging the difficulty of a piece is always subjective, yet it remains a necessary part of teaching. Music teachers depend on such judgments to match repertoire to each student's skills and to plan steady progress.

Automatic estimation of piece difficulty is an interesting task in the field of music information retrieval (MIR), and could help teachers and students find suitable repertoire, support more objective performance evaluations, and even serve as the basis for recommender systems that suggest pieces by difficulty [32, 34, 43]. Previous work on piano difficulty estimation has mostly focused on the structural aspects of the musical score itself, using symbolic



This work is licensed under a Creative Commons Attribution 4.0 International License. *MOCO '26, Montpellier, France*

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2500-5/26/04

<https://doi.org/10.1145/3802842.3802850>

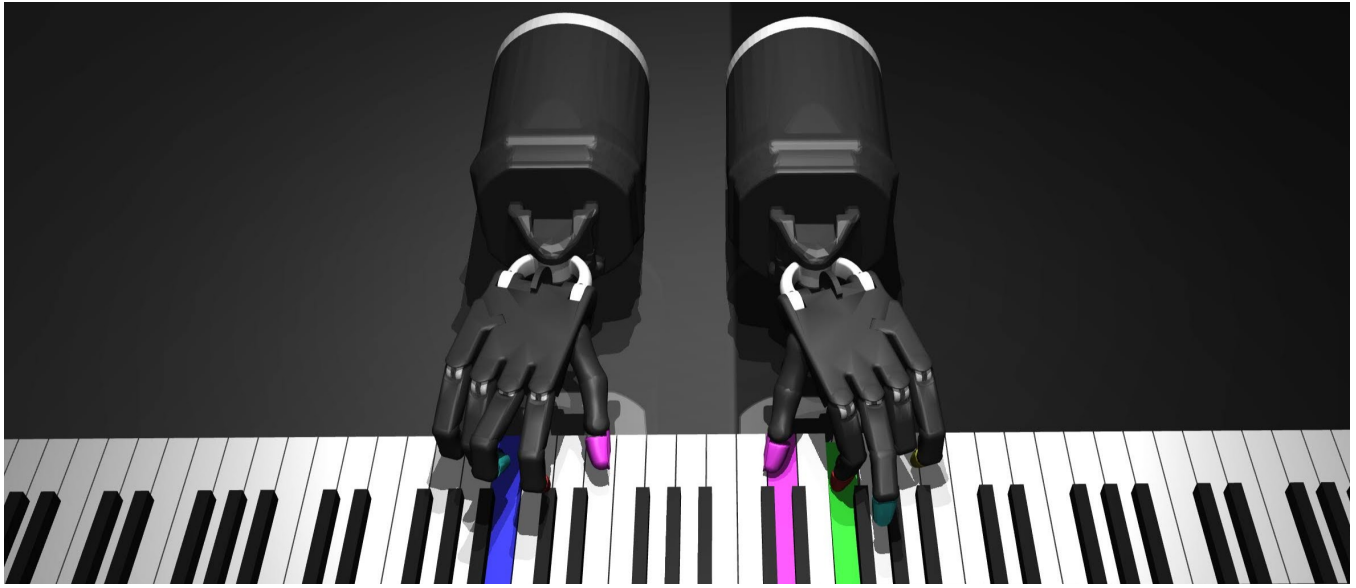


Figure 1: The ROBOPIANIST Simulation in the MuJoCo Physics Engine. See Section 4.1 for details.

features or visual representations to estimate how hard a piece is to play [6, 31].

In this work, we aim to provide an embodied way of assessing difficulty — one that considers how challenging a piece is to learn and perform for a physically simulated pianist. We use reinforcement learning as a framework to relate the agent’s motor control and learning behavior to the difficulty of the task. This shifts the focus from analyzing the musical structure to understanding the effort required by an embodied system to master it. More specifically, we test the hypothesis that an embodied agent’s learning process can serve as a proxy for human-perceived difficulty. We hypothesize that the challenges the agent faces during training — such as complex motor coordination and action planning — are a functional, even if not exact, approximation of the challenges a human learner encounters. This work is intended as a pilot study to explore the feasibility of this *difficulty-through-performance* approach. Our main research question is then: *Can the agent’s learning performance, quantified as how well is the agent able to play a piece of music, serve as an indicator of the difficulty of the piece?*

In this work, we build upon ROBOPIANIST [40], a system of simulated anthropomorphic hands that are able to learn piano pieces using a reinforcement learning approach. The original ROBOPIANIST requires fingering annotations (i.e., an explicit specification of which finger should be used to play each note in the score), and thus it was only able to learn a set of 150 pieces (from the publicly available PIG dataset [23]) for which this information is available. In our approach, we combine ROBOPIANIST with PianoPlayer [22], an automatic fingering generator, thus allowing the system to learn any piece given a symbolic score. Our main goal is to assess the piece’s difficulty by analyzing how well the agent has been able to learn the piece after being given an appropriate training time. The proposed approach would address several of the challenges listed

above: the embodied agent would need to (1) develop motor coordination in both hands; (2) develop rhythmic accuracy in order to perform a piece of music; and it would need to (3) develop a motor program mapping notes in the score to which finger it should use for each note, as well as the corresponding keys in the keyboard. For the sake of simplicity, in this work we don’t consider the aspects of expression and musicality (nor do we consider the visual aspects of “decoding” the visual information in the score). Figure 1 shows an image of the ROBOPIANIST in the simulated environment.

The remainder of this paper is structured as follows: Section 2 reviews related work. Section 3 describes the dataset. Section 4 describes the robotic hands framework used in the experiments. Section 5 discusses estimating piece difficulty by using the learning performance of the robotic hands as a proxy. Section 6 presents the experimental setup. The results of the experiments are discussed in Section 7. Finally, Section 8 concludes this paper.

2 Related Work

In this section, we discuss the literature on robotic music performance and piano piece difficulty estimation.

2.1 Robotic and Embodied Performance

Robotic music performance is challenging because it involves precise coordination and timing. Early approaches focused on specialized robots or pre-programmed sequences to press keys or strings accurately [15, 18, 41], while more recent work looks at dexterous control and coordination across multiple effectors to play more complex pieces [39].

The RoboPianist framework [40] is a large-scale simulation benchmark for piano playing, with tasks ranging from simple note sequences to full polyphonic pieces using realistic piano mechanics and robotic hand models. It offers a way to study motor control, imitation, and reinforcement learning in musical performance.

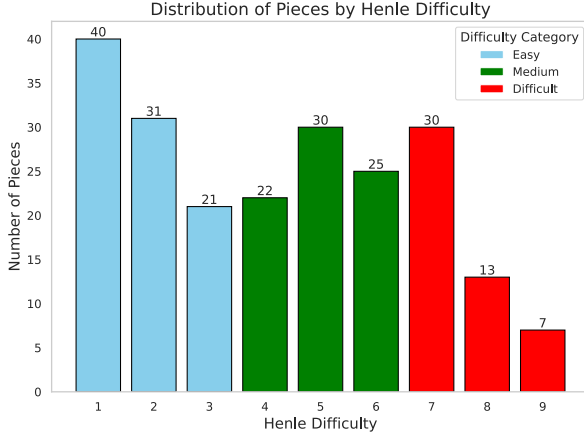


Figure 2: The number of easy, medium, and difficult pieces in the dataset according to the Henle difficulty ratings.

Building on this environment, Qian et al. introduced an imitation learning based framework to train a generalist agent capable of playing unseen piano pieces [28].

2.2 Piano Difficulty Estimation

Prior research has already explored diverse methodologies and datasets to successfully categorize piano sheet music according to its performance difficulty [6, 31, 33, 34].

A regression-based approach using symbolic music features has been proposed by Chiu et al. [6] already in 2012. By using Rreli-eff [33] - a feature selection approach - they introduced new relevant feature sets to capture difficulty characteristics. Their work demonstrated that automatically predicting the difficulty of a piece was theoretically feasible, though it also revealed limitations using existing features. In 2022, Ramoneda et al. [32] incorporated piano fingering algorithms and showed that Gated Recurrent Unit (GRU) neural networks with attention mechanisms over gradient-boosted trees achieved higher interpretability and accuracy. One year later, they [31] also proposed a transformer-based approach to estimate the difficulty of a piano piece from sheet music images. Their method achieved state-of-the-art performance on five datasets. In 2024, Ramoneda et al. introduced the "Can I Play It?" (CIPI) dataset [29], a piano score dataset with annotated difficulty labels from G. Henle Verlag [16]. Their experiments provide a useful foundation for further research. In contrast to symbolic or vision-based approaches, we infer difficulty implicitly from embodied learning behavior and try to bridge the gap between robotic and cognitive models of musical skill.

3 Dataset

In order to assess the difficulty of a piece by analyzing the learning performance of the simulated robotic agent, we need data consisting of musical scores of piano pieces with difficulty labels. For this project, we put together a dataset consisting of 219 piano scores in MIDI format.

While there exist several attempts to categorize piano pieces according to their perceived difficulty [26, 27], one of the most used categorization systems in research is the one introduced by G. Henle Verlag [29, 32, 43]. The difficulty rankings were annotated by Prof. Rolf Koenen [16] and feature nine levels of difficulty, whereas they were categorized into three classes: easy (1-3), medium (4-6), and difficult (7-9).

We obtained the scores from three different sources:

- (1) **The Mikrokosmos-difficulty dataset.** Introduced by Ramoneda et al. [32], this dataset consists of pieces from Bartók's Mikrokosmos collection. In this work, 146 of these pieces are used. The dataset provides symbolic score data in MusicXML format (which we converted to MIDI using the *partitura* library¹) and constitutes the majority of the data used in this work.
- (2) **ASAP Dataset.** Another major part of the data comes from a subset of 47 pieces from the ASAP Dataset (Aligned Scores and Performances) [25]. ASAP contains aligned musical scores (MIDI and MusicXML) and performances (MIDI and audio), with annotations for beats, downbeats, time, and key signatures. From this dataset, we discarded very long pieces (longer than 5 minutes) as well as a small subset of pieces - most notably the Chopin Études - where we observed minor inconsistencies in the musical scores.
- (3) **MuseScore** To address the issues with the Chopin Études in ASAP and to expand the repertoire, we supplemented the dataset with 26 selected pieces from MuseScore,² sourced from the user venson72150 [21]. These scores were manually checked for accuracy, and include primarily the missing Chopin Études (Op. 10 and Op. 25) as well as a few additional pieces. The selected scores are available in both MIDI and MusicXML format, exported directly from the MuseScore platform.

Figure 2 shows the distribution of difficulty across all pieces of the dataset.

4 Methodology

4.1 RoboPianist

The main building block of this project is the ROBOPIANIST system by Zakka et al. [40]. This system uses a software simulation of existing robotic hands, Shadow Dexterous Hand by Shadow Robot³, implemented in the open-source MuJoCo Physics engine [36], as well as a simulation of a full-size standard 88-key digital piano. The Shadow Dexterous Hands are anthropomorphic five-fingered devices with 24 degrees of freedom, designed to closely replicate the kinematics and dexterity of the human hand. Figure 1 depicts the ROBOPIANIST simulation in the open-source MuJoCo Physics engine.

Using MuJoCo, Zakka et al. simulated each key as joints with linear springs, which were regarded as active if their "angles" were within 0.5 degrees of the maximum range. MIDI files were converted into time-indexed note trajectories (i.e., piano rolls), with each note being represented by a one-hot vector of length 88. They then

¹<https://github.com/CPJKU/partitura/>

²<https://musescore.com>

³<https://www.shadowrobot.com/dexterous-hand-series/>

Observations	Unit	Size
Hand and forearm joints	rad	52
Forearm Cartesian position	m	6
Piano key joints	rad	88
Active fingers	discrete	$L \cdot 10$
Piano key goal state	discrete	$L \cdot 88$

Table 1: The observation space for the original ROBOPIANIST, where L represents the lookahead horizon (Table adapted from [40]).

evaluated the pianist’s performance using an F1-metric score, which combines both recall and precision. Recall reflects how accurately the correct “on” keys were pressed, while precision captures how well unintended “off” keys were avoided. A high F1 score, therefore, indicates both accurate key presses and the absence of incorrect notes.

The Reinforcement Learning (RL) approach begins with the formulation of the piano-playing task as a finite-horizon Markov-Decision-Process (MDP), where the goal is the maximization of the total expected (discounted) reward over the horizon by minimizing the difference between pressed and targeted keys. An overview of the ROBOPIANIST observation space can be found in Table 1.

The ROBOPIANIST’s actions are 45-dimensional, consisting of both the target joint angles for the hand and a scalar value for handling the use of the sustain pedal.⁴

A naive approach to piano playing would be to reward the agent simply for pressing the correct keys at the correct times. However, this is insufficient. Playing the right key is not enough—the agent must also plan finger and hand movements to reach subsequent notes efficiently. Pressing a key with a finger or hand configuration that cannot reach the next note in time will cause the pianist to struggle, even if the current note is correct. To address this, the agent must anticipate future notes. This is achieved by stacking goal states over a lookahead horizon L , allowing the agent to plan several steps ahead and move its hand strategically to follow upcoming notes.

This shows that the agent not only needs to plan ahead but also needs to know which finger should press each key. Getting the fingering right is a crucial part of being able to play the piano well.

Being aware of the importance of this subtask, Zakka et al. annotated MIDI files by adding fingering information from the PIG dataset, a collection of classical piano pieces containing fingering annotations for all music notes, developed by Nakamura et al. [23]. For their approach, they assembled 150 annotated MIDI files (the REPERTOIRE-150 dataset) containing piano pieces. The dataset includes pieces of different levels of difficulty and duration, written by 24 composers from the Common Practice Period.⁵

Due to the high dimensionality of the action space, exploration posed a major challenge for the agent. To overcome this, Zakka et

al. incorporated the finger labels into a finger reward term given by Equation 1:

$$r_{\text{finger}} = \left(\frac{1}{K} \sum_i^K g(\|p_f^i - p_k^i\|_2) \right) \quad (1)$$

where g is the tolerance function from `dm_control` [37] that maps the L2 distance between the Cartesian positions of the finger $p_f \in \mathbb{R}_3$ and the corresponding key $p_k \in \mathbb{R}_3$ to a bounded reward value in $[0,1]$.

The key reward incorporates a reward for pressed target keys and a penalty when false keys are pressed strongly enough to create a sound, as shown in Equation 2. Zakka et al. excluded the relevance of the number of false positives that are pressed to avoid causing the agent to be too conservative and not press any key at all.

$$r_{\text{key}} = \frac{1}{2} \left(\frac{1}{K} \sum_i^K g(\|k_s^i - 1\|_2) \right) + \frac{1}{2} (1 - 1_{\{\text{false positive}\}}) \quad (2)$$

Here, K is the number of keys that should be pressed at the given timestep. The normalized key joint position is specified by $k_s \in (0, 1)$ and $1_{\{\text{false positive}\}}$ is an indicator equal to 1 if any unintended key produces sound.

In addition to the key and finger reward terms, they also included an energy minimization term to evade erratic and pointless movements and to reduce variance in the agent’s behavior. The energy expenditure term, formed by the dot product of the magnitudes of the joint torque vector and the joint velocity vector, is given by Equation 3:

$$r_{\text{energy}} = |\tau_{\text{joints}}|^T |v_{\text{joints}}| \quad (3)$$

Putting everything together and assigning weights to the individual reward components, we obtain the total reward function to train the agent to play the piano, as expressed in Equation 4.

$$r_{\text{total}} = r_{\text{key}} + r_{\text{finger}} - 0.005 \cdot r_{\text{energy}} \quad (4)$$

The authors then restricted the action space by disabling certain inessential degrees of freedom of the robot’s hands and reducing the thumb joint range. Although they reached similar results without doing this, it led to a significant speedup in the learning process.

Zakka et al. used Dropout-Q-Learning (DroQ) [13], a state-of-the-art model-free RL algorithm, to teach the agent a useful policy. DroQ is a variation of the Randomized ensemble double Q-learning (REDQ) [5] and the widely known Soft-Actor-Critic (SAC) algorithm [12], which was used by Xu et al. in their piano-playing approach. One of DroQ’s core advantages against REDQ lies in its superior computational efficiency, which is achieved by using only a small ensemble of Q-functions with dropout connection and layer normalization. Dropout-Q-Learning is double efficient - i.e., computationally and sample efficient.

4.2 Fingering Prediction Integration

Piano sheet music typically only specifies the keys to be played, but not which finger should be used to press them. Yet for playing the same piano passage, there are usually numerous possible finger combinations, and the choice for each key deeply influences the subsequent finger selections.

⁴Instead of using all 24 degrees of freedom for each hand, Zakka et al. restricted the motion of the Shadow Dexterous Hands, resulting in 22 degrees of freedom for each hand, plus the scalar value for the pedal, resulting in this 45-dimensional action space [40].

⁵The Common Practice Period refers to the music from the 18th to the early 20th centuries, and corresponds to music from the Baroque to the late Romantic period.

Fingering prediction refers to the automatic assignment of piano fingers (1–5) to individual notes in a score. The optimal choice depends strongly on factors such as the pianist’s hand size, hand position, and, to some extent, the pianist’s technical skill level. Fingering prediction has been studied from both a musicological perspective and in the context of computational modeling (e.g., [23, 30]).

We do not aim to develop new fingering prediction algorithms for this project, as it is beyond the scope of our work. Instead, we rely on existing tools to provide fingering predictions. Among those existing tools, the open-source PianoPlayer project by Marco Musy [22] was chosen as it is one of the most widely used automatic fingering generators available.

The algorithm tries to minimize the necessary finger speed to play a piano score by searching through possible fingering combinations. In contrast to simple look-up tables, this process is completely dynamic as it considers the moving positions, the fingers’ speed, and the notes’ durations [22].

The algorithm’s goal is to reduce unnecessary hand movements as effectively as possible. As optimal fingerings also depend on the size of a player’s hand, the algorithm provides the option of specifying seven predefined hand sizes for the annotations.

We adapted ROBOPIANIST’s preprocessing function in order to first predict an appropriate fingering with PianoPlayer, thus allowing the agent to learn any piece of music.

4.3 Reinforcement Learning Setup

ROBOPIANIST is trained using Dropout Q-learning (DroQ), a regularized variant of the Soft-Actor-Critic (SAC) algorithm [12].

As in standard deep Q-learning, past states and transitions are stored in a replay buffer, enabling the agent to revisit prior experience and improve stability during policy updates [20]. For ROBOPIANIST, these states consist of the joint positions and velocities for both Shadow Dexterous Hands, the Cartesian positions of all fingertips, and the current state of all 88 piano keys, as well as the sustain pedal. Additionally, they include the target goal states - i.e., which keys and/or the pedal should be pressed - for the current timestep and for a short prediction horizon L (typically around 10 future steps). In total, this results in a state vector of approximately 1.1 k dimensions, depending on configuration.

The agent consists of multiple neural networks: the actor is a single 3-layer multilayer perceptron (MLP), while both the critic and target critic are ensembles of two MLPs each. The actor and critic share the same architecture, with hidden layers of 256 neurons. In the critic networks, each linear layer is followed by dropout (rate 0.01) and layer normalization. The actor uses a tanh-diagonal-Gaussian activation, whereas the critics employ ReLU activations. A separate temperature network automatically adjusts the entropy coefficient, balancing exploration and exploitation during training. The architecture of the agent is illustrated in Figure 3.

Training is performed for a number of environment steps according to the length of the specific piano piece using a batch size of 256, a discount factor $\gamma = 0.82$, and a control time step of 0.05 s. The F1 score performance is evaluated every 10,000 training steps. For all networks, we used standard Adam optimizers with learning rates of 3×10^{-4} .

5 Difficulty Estimation via Learning Performance

After training on a piece of music, the performance of the agent reaches a certain level of accuracy, which can be measured by the F1 score. The F1 scores on easier pieces will, of course, be higher than those on more difficult ones, where it is harder for the agent to achieve the same level of performance. Thus, we can expect an inverse proportionality between the reached F1 score after training and the song’s inherent difficulty level. We assess the difficulty of a piece in the same manner as it was done by Prof. Rolf Koenen for G. Henle Verlag, using 9 grades, whereas higher grades correspond to higher levels of difficulty.

Depending on the pieces’ difficulties, the ROBOPIANIST reaches an F1 score of up to around 0.9 for easier pieces, but for more difficult pieces, it usually does not exceed a score of approximately 0.3. Since longer pieces naturally require more time (i.e., a higher number of training epochs) to master than shorter pieces, we must consider piece lengths when determining the number of epochs as well.

In practice, the number of training epochs was selected based on a rough heuristic informed by a few initial runs. Early experiments indicated that a training duration on the order of $\frac{4T}{60} \cdot 10^6$ epochs (for a piece of duration T) was generally sufficient to produce reasonably stable performance levels. This value should be regarded as a rough estimate informed by prior observations rather than a rigorously determined quantity. However, further testing suggested that training for fewer epochs was sufficient to obtain reliable difficulty estimates. To reflect this and to accommodate time constraints, the final number of epochs $E(T)$ was chosen as a scaled-down fraction of the initial estimate.

As in most DRL applications, training is computationally intensive, especially under hardware limitations. Therefore, our goal is to minimize training duration while maintaining acceptable accuracy. To this end, we introduce a speed factor ν , which adjusts the number of epochs according to the desired trade-off between training time and accuracy. It is important to note that this factor refers exclusively to computational training duration and has no relation to the musical tempo at which pieces are played. Intuitively, a smaller ν -value corresponds to longer training and higher accuracy, while a larger ν yields shorter training at the cost of reduced precision. In our current setup, the speed factor can take any integer value from 1 to 3.

We set the number of training epochs E depending on the time T of the piano piece and the speed factor ν in the following manner, as presented in Equation 5:

$$E_\nu(T) = \frac{T}{60} \cdot 2^{-\nu} \cdot 10^6 \quad | \quad \nu \in \{1, 2, 3\} \quad (5)$$

Assuming an inverse proportionality between the piece’s difficulty level and the achieved F1 score after training, we can model the difficulty level as a logarithmic regression problem described by Equation 6:

$$G = k_{\nu 1} \log(f_1 k_{\nu 2}) \quad (6)$$

whereas G represents the difficulty grade, f_1 the F1 performance score after training for E epochs and $k_{\nu 1}$, $k_{\nu 2}$ are factors which have to be determined empirically, specific to each speed factor.

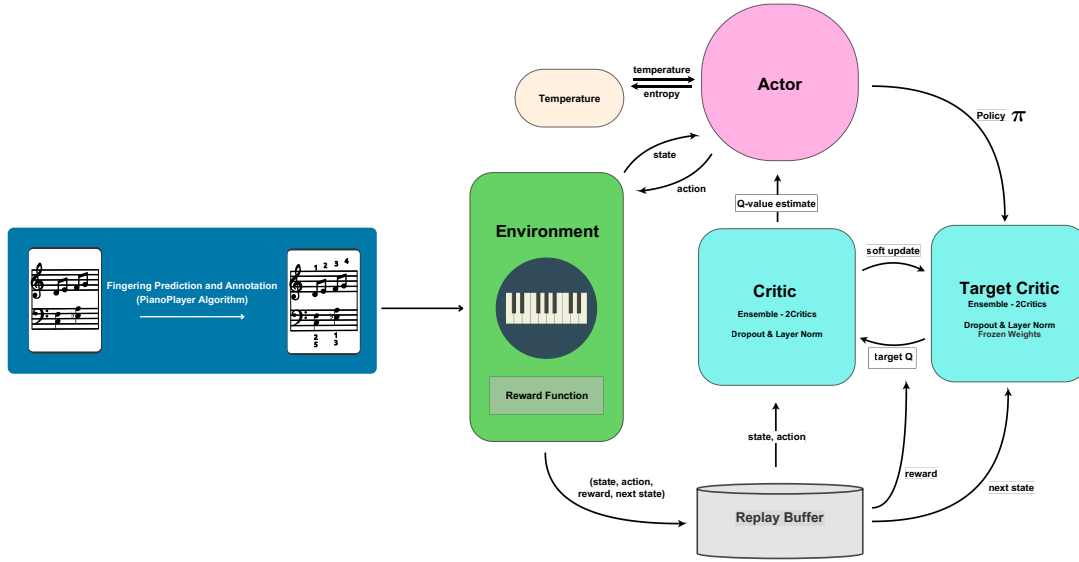


Figure 3: Architecture and Reinforcement Learning Setup for the ROBOPIANIST approach

6 Experimental Setup

In order to assess how well the learning performance of the ROBOPIANIST framework indicates piece difficulty, we conduct two experiments:

- (1) **Experiment 1:** We test whether there is a statistical relationship between the learning performance of the ROBOPIANIST, measured by its F1 scores, and piece difficulty.
- (2) **Experiment 2:** We investigate whether we can use the learning performance of the ROBOPIANIST in a classification context to estimate the difficulty of an unseen piece.

6.1 Experiment 1

To investigate the correlation of piece difficulty and the performance of the RoboPianist, as measured by the F1 score and modeled by the logarithmic regression described in Equation 6 in Section 5, we performed a linear regression analysis [38] for each of the speed factors ($\nu = 1, 2, 3$). In this setup, the dependent variable is the difficulty G , and the independent variable is the $\log(\text{F1})$ score.

6.2 Experiment 2

For the second experiment, we use a classification setup to test the accuracy of classifiers predicting difficulty for unknown pieces using the learning performance of the ROBOPIANIST framework.

We split our dataset of 219 pieces into training (80%) and test (20%) sets using a stratified split to keep the class distribution roughly the same. The test set contained 44 pieces. As a baseline, we considered a majority-class predictor that always chooses the most common class in the test set, which would be correct about 18% of the time. We can use this baseline as reference point to evaluate whether our classifiers are actually learning meaningful patterns beyond class frequency. By using a simple Support Vector Machine (SVM), we achieve high accuracies. However, the dataset

is imbalanced as we have a higher number of easy pieces and fewer difficult ones. As we can see in Figure 2, it is easier to differentiate between pieces with lower Henle levels. Thus, we have to account for this imbalance of the dataset if we want to fairly evaluate the classifiers' performance.

We employ the Balanced Accuracy metric [3], calculating the average recall across all classes, regardless of the number of samples per class. This way, all classes are treated equally. By doing this and by allowing a tolerance of ± 1 , we reach a relaxed balanced accuracy of up to 81.5% using the lowest speed factor. For higher speed factors, however, the accuracy values decrease substantially, as we can see in Table 3.

Furthermore, given the highly subjective nature of the difficulty labels,⁶ we also consider a relaxed accuracy scenario, i.e., allowing predictions to be within ± 1 of the *annotated* difficulty level.

7 Results and Discussion

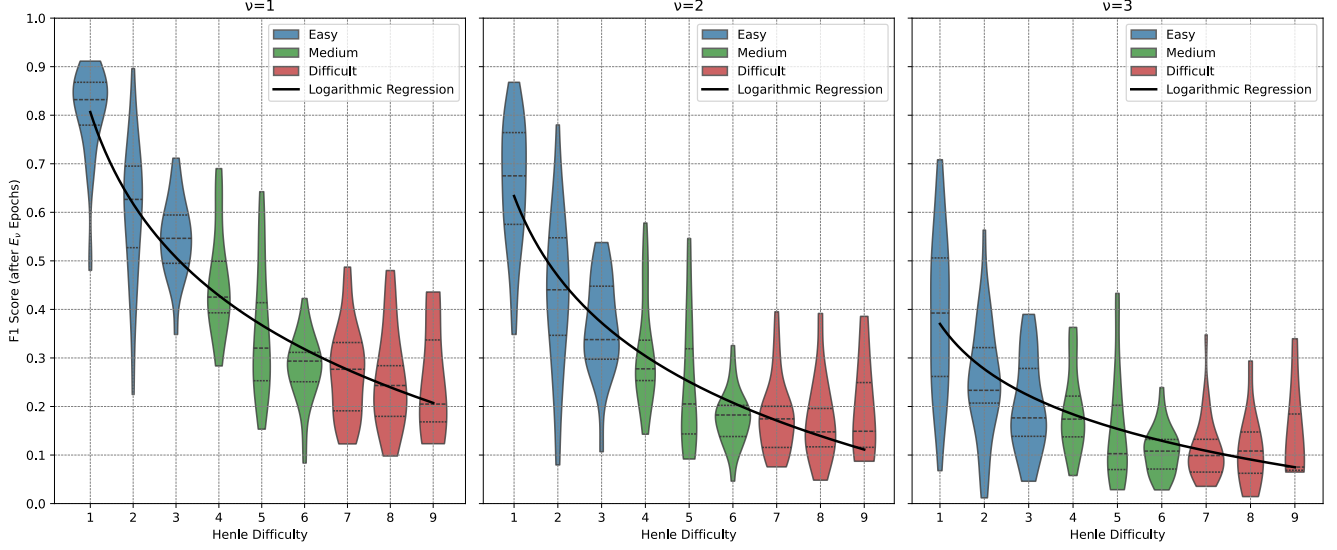
7.1 Experiment 1

After training the agent for E_ν epochs on a set of 219 pieces of piano music, we can see the following relations:

The plots in Figure 4 reveal an inverse relation between the median F1 scores and Henle difficulty levels of the pieces. At higher speed factors, however, this relationship becomes less pronounced, making it more challenging to determine the difficulty levels based on the F1 scores.

Using the previously suggested approach and regarding the difficulty level as a logarithmic regression problem (see Equation 6), we can approximate the relationship as depicted in Figure 4.

⁶As Rolf Koenen points out: "Any evaluation of art or music will always be subjective, even if the aim was to be objective. Despite the fact that I have endeavoured to be as careful as possible, I am all too aware that the results of my work can be called into question, [...]" [16].

F1 Scores by Henle Difficulty**Figure 4:** Violin plots showing the distribution of F1 scores across the nine Henle difficulty levels for each speed factor ($v = 1, 2, 3$).

Statistic	$v = 1$	$v = 2$	$v = 3$
R^2 (Effect Size)	0.664	0.556	0.344
t -statistic ($\log(F1)$)	-20.694	-16.487	-10.661
p -value ($\log(F1)$)	< 0.001	< 0.001	< 0.001

Table 2: Linear Regression Analysis results. The t -statistic and p -value confirm the significance of the $\log(F1)$ score as a predictor of the Henle Grade (G).

Metrics	Baseline	$v = 1$	$v = 2$	$v = 3$
Imbalanced Accuracy	0.182	0.568	0.432	0.318
Balanced Accuracy	0.111	0.486	0.357	0.261
Relaxed Imbalanced Accuracy (± 1)	0.318	0.909	0.727	0.682
Relaxed Balanced Accuracy (± 1)	0.222	0.815	0.671	0.632

Table 3: Classification results for different speed factors ($v = 1, 2, 3$) using the ROBOPIANIST performance data, showing balanced and imbalanced accuracy metrics. We focus on relaxed balanced accuracy (± 1).

To statistically validate the inverse relationship observed in Figure 4, Table 2 summarizes the results of the linear regression analyses for three training speed factors ($v = 1, 2, 3$). Across all speed factors, the results show a significant effect ($p < 0.001$) of F1 score being a strong predictor of piece difficulty for all speed factors. The effect size, measured by R^2 , decreased as training time was reduced, showing that performance becomes less predictable with shorter training.

Overall, the results confirm a clear inverse correlation between F1 score and Henle grade: harder pieces tend to produce lower F1 scores. This shows that F1 score can be used as a reliable indicator of piece difficulty.

7.2 Experiment 2

As we can see in the plots shown in Figure 4, there is a lot of variation in the reached F1 scores for pieces within the same Henle difficulty level.

Given that the relations between F1 scores and difficulty levels strongly depend on the number of training epochs, and consequently on the speed factor, we need to train three individual

classifiers to predict the difficulty in each situation. Table 3 reports the results of the classification experiment.

To assess whether the observed differences in classification accuracy were statistically significant, we conducted pairwise comparisons using a binomial test. Each classifier trained at a different speed factor ($v = 1, 2, 3$) was compared both to the majority-class baseline and to each other.

As shown in Table 4, classifiers trained with longer training times (lower v) significantly outperformed the majority-class baseline ($p < 0.001$ for $v = 1, 2$; $p = 0.029$ for $v = 3$). Among the trained models, the difference between $v = 1$ and $v = 3$ was also significant ($p = 0.001$), while other pairwise differences did not reach significance ($p > 0.05$).

While not directly comparable, the classification accuracies obtained here are similar to those reported in other approaches for piece difficulty estimation focusing on musical structure from the score [32, 43]. It is important to note, however, that the approach presented here is not meant to be a computationally efficient method, or to beat the state-of-the-art. Instead, it aims to show that we can use the learning performance of an embodied agent trained by

	$\nu = 1$	$\nu = 2$	$\nu = 3$
Baseline	< 0.001	< 0.001	0.029
$\nu = 1$	—	0.093	0.001
$\nu = 2$	—	—	0.170

Table 4: Pairwise p -values from binomial tests comparing classification accuracies across speed factors (ν) and against the majority-class baseline.

trying to play the right notes while having efficient movements (modeled by the r_{energy} term in the reward function; see Equation 4) is viable framework to study music performance related questions.

8 Conclusions and Future Directions

In this work we presented an embodied approach to estimating the difficulty of a piano piece using an embodied approach by considering the performance of robotic hands.

The classification results presented in Section 6 show that our assessment framework seems to achieve similar performance to state-of-the-art difficulty classifiers, even though a direct comparison is complicated since we have used a mixture of different datasets for our experiments [29, 31]. Nevertheless, we also have to consider time factors to fairly compare performances. Unlike other algorithms that offer faster results, our approach requires additional computation, as our agent has to be trained on the pieces we want to classify first. However, that being said, when prioritizing accuracy over speed, our method remains highly competitive and makes it possible to assess the difficulty of specific parts of a piece. While our algorithm might perform better - after being trained on more data - further testing on datasets like CIPI and others would be necessary to comprehensively compare it to other state-of-the-art classifiers.

Considering that - due to time restrictions - the classifier has so far only been trained on a rather small dataset of 219 pieces, we can expect its performance to substantially increase as the dataset grows. Adding the CIPI dataset by Ramoneda et al. [29] could be particularly valuable in improving our model, as could any additional piano piece assessed by Ralf Koenen for the Henle Verlag, considering his relatively consistent evaluation method. Any further dataset extensions might lead to significant improvements and open up new future research directions.

Even though assessing the difficulty of a piece of music remains an inherently subjective task, our proposed algorithm enables a rather consistent grading scheme. This makes it easier to compare the difficulty of new pieces of piano music with those already assessed.

On a technical side, a bottleneck for training the model is the choice of MuJoCo as the physics engine. Many existing physics engines, while effective, can be too slow for complex tasks like robotics training, particularly when dealing with high-dimensional data or long simulation runs. The high computational cost creates a significant bottleneck, limiting the scope of what researchers can explore.

To overcome this, a promising path is to adopt new, faster physics engines. For example, the recently introduced Genesis physics engine [2] is reported to offer massively accelerated speeds and high physical fidelity, making it possible to run experiments that were previously unfeasible. A possible future research direction would be to transfer the whole framework to such an engine, which might drastically reduce training times, and could enable both very fast and accurate difficulty assessment as well as allow for more ambitious research in the domain of highly complex robotics.

The current work only focused on hands, but pianists are not disembodied hands, but full-bodied systems. Future work will expand this work by combining the Shadow Dexterous Hands with a full-body humanoid robot, allowing for exploring how body motion is developed. Furthermore, we aim to experiment with more sophisticated piano fingering algorithms, including learning models that both learn a fingering and the movements from scratch, in a way more similar to what human pianists do.

While human pianists typically rely on varying practice tempi - starting at a slower pace before gradually increasing speed - our agent currently rehearses and performs at a fixed, static tempo without musical expressivity. Experimenting with flexible practice tempi would be an intriguing research direction for future work. Additionally, given that robust computational models for rendering expressive performances already exist [4, 14, 42], integrating these models into our framework would allow us to move beyond constant-tempo, static renditions.

Finally, we are interested in exploring the use of such an embodied reinforcement learning agent to explore and assess the quality of movements by human pianists. Such a system could be used for e.g., providing real-time feedback in the context of piano practice.

Acknowledgments

This work was supported by the Austrian Science Fund (FWF) under grants and PIN 1347924 (“AURA: Augmenting musical interaction via EVAs”) and PAT 8820923 (“Rach3: A Computational Approach to Study Piano Rehearsal”).

References

- [1] ABRSM. [n. d.]. *About Piano exams*. <https://www.abrsm.org/en-gb/instruments/piano> (Accessed: 2025-10-02).
- [2] Genesis Authors. 2024. Genesis: A Generative and Universal Physics Engine for Robotics and Beyond. <https://github.com/Genesis-Embodied-AI/Genesis> (Accessed: 2026-03-30).
- [3] Kay Henning Brodersen, Cheng Soon Ong, Klaas Enno Stephan, and Joachim M. Buhmann. 2010. The Balanced Accuracy and Its Posterior Distribution. In *2010 20th International Conference on Pattern Recognition*. Istanbul, Turkey, 3121–3124. doi:10.1109/ICPR.2010.764
- [4] Carlos E. Cancino-Chacón, Maarten Grachten, Werner Goebel, and Gerhard Widmer. 2018. Computational Models of Expressive Music Performance: A Comprehensive and Critical Review. *Frontiers in Digital Humanities* Volume 5 (2018), 25. doi:10.3389/fdigh.2018.00025
- [5] Xinyue Chen, Che Wang, Zijian Zhou, and Keith W. Ross. 2021. Randomized Ensembled Double Q-Learning: Learning Fast Without a Model. In *International Conference on Learning Representations (ICLR)*. Vienna, Austria. <https://openreview.net/forum?id=AY8zfZm0tDd>
- [6] Shih-Chuan Chiu and Min-Syan Chen. 2012. A Study on Difficulty Level Recognition of Piano Sheet Music. In *2012 IEEE International Symposium on Multimedia*. Irvine, California, USA, 17–23. doi:10.1109/ISM.2012.11
- [7] Emily G. Eldridge. 2025. Piano at the Center: Reimagining Music Education. *Journal of General Music Education* 38, 2 (2025), 26–32. doi:10.1177/27527646241300615
- [8] K Anders Ericsson, Ralf T Krampe, and Clemens Tesch-Römer. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 100, 3 (1993), 363.

- [9] Rolf Inge Godøy and Marc Leman (Eds.). 2009. *Musical Gestures: Sound, Movement, and Meaning* (1 ed.). Routledge, New York, NY, USA.
- [10] Werner Goebel and Caroline Palmer. 2008. Tactile Feedback and Timing Accuracy in Piano Performance. *Experimental Brain Research* 186, 3 (April 2008), 471–479. doi:10.1007/s00221-007-1252-1
- [11] Werner Goebel and Caroline Palmer. 2013. Temporal control and hand movement efficiency in skilled music performance. *PLOS One* 8, 1 (2013), e50901.
- [12] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. 2018. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*. Stockholm, Sweden. arXiv:1801.01290 [cs.LG]
- [13] Takuya Hiraoka, Takahisa Imagawa, Taisei Hashimoto, Takashi Onishi, and Yoshimasa Tsuruoka. 2022. Dropout Q-Functions for Doubly Efficient Reinforcement Learning. In *International Conference on Learning Representations (ICLR)*. arXiv:2110.02034 [cs.LG]
- [14] Dasaem Jeong, Taegyun Kwon, Yoojin Kim, Kyogu Lee, and Juhan Nam. 2019. VirtuosoNet: A Hierarchical RNN-based System for Modeling Expressive Piano Performance. In *International Society for Music Information Retrieval Conference (ISMIR)*. 908–915.
- [15] Ichiro Kato, Sadamu Ohteru, Katsuhiko Shirai, Toshiaki Matsushima, Seinosuke Narita, Shigeki Sugano, Tetsunori Kobayashi, and Eizo Fujisawa. 1987. The robot musician 'wabot-2' (waseda robot-2). *Robotics* 3, 2 (June 1987), 143–155. doi:10.1016/0167-8493(87)90002-7
- [16] R. Koenen. [n. d.]. *Prof. Rolf Koenen notes about Henle Verlag difficulty classification*. <https://www.henle.de/en/Levels-of-Difficulty/> (Accessed: 2025-01-14).
- [17] Alexander Lerch, Claire Arthur, Ashis Pati, and Siddharth Gururani. 2020. An Interdisciplinary Review of Music Performance Analysis. *Transactions of the International Society for Music Information Retrieval* (Nov 2020). doi:10.5334/tismir.53
- [18] Jen-Chang Lin, Hsing-Hsin Huang, Yen-Fang Li, Jen-Chao Tai, and Li-Wei Liu. 2010. Electronic piano playing robot. In *2010 International Symposium on Computer, Communication, Control and Automation (3CA)*, Vol. 2. 353–356. doi:10.1109/3CA.2010.5533457
- [19] Christine L MacKenzie and DL Van Eerd. 1990. Rhythmic precision in the performance of piano scales: Motor psychophysics and motor programming. In *Attention and Performance: Motor Representation and Control*. Hillsdale, NJ.: Erlbaum, 375–408.
- [20] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fiedelnd, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (2015), 529–533.
- [21] MuseScore. [n. d.]. *verson72150*. <https://musescore.com/user/33306646> (Accessed: 2025-01-15).
- [22] Marco Musy. 2021. PianoPlayer: Automatic piano fingering generator. <https://github.com/marcomusy/pianoplayer> (Accessed: 2025-01-14).
- [23] Eita Nakamura, Yasuyuki Saito, and Kazuyoshi Yoshii. 2020. Statistical learning and estimation of piano fingering. In *Information Sciences*, Vol. 517. Kyoto, Japan, 68–85. doi:10.1016/j.ins.2019.12.068
- [24] Caroline Palmer. 1997. Music performance. *Annual Review of Psychology* 48, 1 (1997), 115–138.
- [25] Silvan David Peter, Carlos Eduardo Cancino-Chacón, Francesco Foscarin, Andrew Philip McLeod, Florian Henkel, Emmanouil Karystinaios, and Gerhard Widmer. 2023. Automatic Note-Level Score-to-Performance Alignments in the ASAP Dataset. *Transactions of the International Society for Music Information Retrieval* 6, 1 (June 2023), 27–42. doi:10.5334/tismir.149
- [26] Piano Library. [n. d.]. *A Thematic Catalogue of the Solo Piano Repertoire*. <https://www.pianolibrary.org/> (Accessed: 2025-01-15).
- [27] Piano Syllabus. [n. d.]. *Music Database*. <https://pianosyllabus.com/> (Accessed: 2025-01-15).
- [28] Cheng Qian, Julien Urain, Kevin Zakka, and Jan Peters. 2024. PianoMime: Learning a Generalist, Dexterous Piano Player from Internet Demonstrations. In *8th Annual Conference on Robot Learning*. <https://openreview.net/forum?id=t0LkF9JnVb>
- [29] Pedro Ramoneda, Dasaem Jeong, Vsevolod Eremenko, Nazif Can Tamer, Marius Miron, and Xavier Serra. 2024. Combining Piano Performance Dimensions for Score Difficulty Classification. *Expert Systems with Applications* 238 (2024), 121776. doi:10.1016/j.eswa.2023.121776
- [30] Pedro Ramoneda, Dasaem Jeong, Eita Nakamura, Xavier Serra, and Marius Miron. 2022. Automatic Piano Fingering from Partially Annotated Scores using Autoregressive Neural Networks. In *Proceedings of the 30th ACM International Conference on Multimedia (Lisboa, Portugal) (MM '22)*. Association for Computing Machinery, New York, NY, USA, 6502–6510. doi:10.1145/3503161.3548372
- [31] Pedro Ramoneda, Dasaem Jeong, Jose J Valero-Mas, and Xavier Serra. 2023. Predicting Performance Difficulty From Piano Sheet Music Images. In *Proceedings of the 24th International Society for Music Information Retrieval Conference (ISMIR2023)*. Milan, Italy.
- [32] Pedro Ramoneda, Nazif Can Tamer, Vsevolod Eremenko, Marius Miron, and Xavier Serra. 2022. Score difficulty analysis for piano performance education. In *Proceedings of ICASSP 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- [33] Marko Robnik-Sikonja and Igor Kononenko. 2003. Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning* 53 (2003), 23–69. doi:10.1023/A:1025667309714
- [34] Véronique Sébastien, Henri Ralambondrainy, Olivier Sébastien, and Noël Conrui. 2012. Score Analyzer: automatically determining scores difficulty level for instrumental e-Learning. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*. Porto, Portugal, 571–576. <https://hal.univ-reunion.fr/hal-01187227> C-ACTI.
- [35] John A Sloboda, Jane W Davidson, Michael JA Howe, and Derek G Moore. 1996. The role of practice in the development of performing musicians. *British Journal of Psychology* 87, 2 (1996), 287–309.
- [36] Emanuel Todorov, Tom Erez, and Yuval Tassa. 2012. MuJoCo: A physics engine for model-based control. In *International Conference on Intelligent Robots and Systems (IEEE/RSJ)*. Vilamoura-Algarve, Portugal, 5026–5033. doi:10.1109/IROS.2012.6386109
- [37] Saran Tunyasuvunakool, Alistair Muldal, Yotam Doron, Siqi Liu, Steven Bohez, Josh Merel, Tom Erez, Timothy Lillicrap, Nicolas Heess, and Yuval Tassa. 2020. dm_control: Software and tasks for continuous control. *Software Impacts* 6 (Nov. 2020), 100022. doi:10.1016/j.simpa.2020.100022
- [38] Larry Wasserman. 2004. *All of Statistics: A Concise Course in Statistical Inference*. Vol. 26. Springer, New York, USA.
- [39] Huazhe Xu, Yuping Luo, Shaoxiong Wang, Trevor Darrell, and Roberto Calandra. 2022. Towards Learning to Play Piano with Dexterous Hands and Touch. In *International Conference on Intelligent Robots and Systems (IEEE/RSJ)*. Kyoto, Japan. arXiv:2106.02040 [cs.RO]
- [40] Kevin Zakka, Philipp Wu, Laura Smith, Nimrod Gileadi, Taylor Howell, Xue Bin Peng, Sumeet Singh, Yuval Tassa, Pete Florence, Andy Zeng, and Pieter Abbeel. 2023. RoboPianist: Dexterous Piano Playing with Deep Reinforcement Learning. In *Conference on Robot Learning (CoRL)*. Atlanta, USA.
- [41] Ada Zhang, Mark Malhotra, and Yoky Matsuoka. 2011. Musical piano performance by the ACT Hand. In *2011 IEEE International Conference on Robotics and Automation*. 3536–3541. doi:10.1109/ICRA.2011.5980342
- [42] Huan Zhang, Shreyan Chowdhury, Carlos Eduardo Cancino-Chacón, Jinhua Liang, Simon Dixon, and Gerhard Widmer. 2024. Dexter: Learning and controlling performance expression with diffusion models. *Applied Sciences* 14, 15 (2024), 6543.
- [43] Huan Zhang, Emmanouil Karystinaios, Simon Dixon, Gerhard Widmer, and Carlos Eduardo Cancino-Chacón. 2023. Symbolic Music Representations for Classification Tasks: A Systematic Evaluation. In *Proceedings of the 24th International Society for Music Information Retrieval Conference (ISMIR)*. Milan, Italy.