NON-CONFIDENTIAL

Modeling Player-Specific Behaviors in Elite Chess

Research Project (PRe)

Specialty: Information and Communication Sciences and Technologies (STIC)

Loris SOGLIUZZO

Host Institution:
UCLouvain, ICTEAM, INGI
Place Sainte-Barbe, 2
Louvain-la-Neuve, Belgium

ENSTA Paris Academic Supervisor:
Zhi YAN

Host Institution Supervisor:
Eric PIETTE

Confidentiality Note

This document is non-confidential. It may therefore be consulted online by everyone.

Acknowledgments

I express my sincere gratitude to my supervisor, Prof. Éric Piette, for his trust and guidance throughout these five months. I am deeply grateful for the time he dedicated to mentoring me, for introducing me to the world of academic research, and for providing the opportunity to present my work during a COST Action GameTable meeting.

I am equally grateful to Aloïs Rautureau for his continuous support throughout the project and for our insightful scientific discussions.

I thank both Éric and Aloïs for their meticulous reviewing, valuable feedback, and dedication during the writing and submission of our research paper, "Toward Modeling Player-Specific Chess Behaviors", to the IEEE Conference on Games (CoG). Experiencing the scientific publication process alongside them was an invaluable learning experience.

Finally, I extend my thanks to the ICTEAM laboratory for welcoming me and providing an excellent research environment.

Abstract

The development of chess models has historically prioritized superhuman strength, often producing uninterpretable strategies that diverge from human decision-making. While recent human-centric models, such as Maia-2 and Maia4All, successfully capture behavior at population and individual levels, they primarily target amateur to intermediate players, leaving the modeling of elite world-class players as a significant challenge. Furthermore, current evaluations heavily rely on top-1 move prediction accuracy, a metric that fails to account for the natural diversity and stochasticity inherent in human play. This research extends individual behavior modeling to a cohort of 14 elite chess players active in the 20th century, encompassing world champions and top-tier candidates, while developing evaluation methodologies that respect strategic variance. Building upon the unified architecture of Maia-2, the proposed behavior generation pipeline employs parameter-efficient techniques, specifically player-specific embedding fine-tuning and a Mixture of Experts (MoE) framework with Low-Rank Adaptation (LoRA). To mitigate tactical loop vulnerabilities while preserving human alignment, lightweight Monte Carlo Tree Search (MCTS) and Descent Search, constrained by Nucleus Pruning, are integrated. Finally, to move beyond deterministic metrics like move accuracy and Centipawn error, this work introduces a distributional evaluation pipeline across shared board positions (JSD_{action}) and a stylistic manifold framework (JSD_{spatial}) leveraging AutoEncoders and UMAP dimensionality reduction to compress 2304-dimensional board transitions into a 2D spatial grid.

Keywords: Human-AI Alignment, Behavioral Modeling, Chess Artificial Intelligence, Parameter-Efficient Fine-Tuning, Mixture of Experts (MoE), Stylometry, Jensen-Shannon Divergence.

Contents

List of Figures

List of Tables

Context of the Research Project

This research project was conducted at the ICTEAM Institute (*Institute of Information and Communication Technologies, Electronics and Applied Mathematics*) of the UCLouvain (Louvain-la-Neuve, Belgium) within the Department of Computing Science Engineering (INGI). The research was carried out under the direct supervision of Prof. Éric Piette alongside Aloïs Rautureau, while Prof. Zhi Yan acted as the academic supervisor and graduation committee representative on behalf of ENSTA Paris.

Institutional Framework and Scientific Scope This work is affiliated with the European research initiative COST Action CA22145, GameTable (*A European Network for Artificial Intelligence in Tabletop Games*) (Piette et al., 2024; Rautureau et al., 2026b). Situated at the intersection of behavioral artificial intelligence, game stylometry, and human-AI alignment, the project investigates methods to transition neural policy architectures from pure game-theoretic optimization toward reproducing the nuanced, intuitive decision-making signatures of human players.

Chronology and Research Trajectory The five-month research progression followed four main phases. The work initiated with an in-depth literature review and replication of foundational human-centric chess models, including Maia, Maia-2, and Maia4All. The core development phase focused on architectural design, formulating the parameter-efficient Mixture-of-Experts LoRA (MoE-LoRA) framework and coupling it with tactical search mechanisms, specifically Monte Carlo Tree Search and Descent Search regularized by Nucleus Pruning. The resulting evaluation methodologies on shared board positions led to the research paper entitled *"Toward Modeling Player-Specific Chess Behaviors"* (Sogliuzzo et al., 2026), which was presented during a COST Action GameTable meeting and submitted to the IEEE Conference on Games (CoG 2026). The project concluded with a systematic diagnostic of spatial manifold evaluation biases and the compilation of the complete experimental benchmark.

1. Introduction and Related Work

Historically, chess has served as a benchmark to develop algorithms capable of playing at a superhuman level. This was demonstrated by Campbell et al. (2002) with Deep Blue during the match against Garry Kasparov in 1997. More recently, AlphaZero was introduced by Silver et al. (2017), demonstrating that a competitive model can be trained without reliance on human game records, guided solely by explicit game rules and domain-specific convolutional inductive biases through reinforcement learning from self-play.

However, superhuman engines are optimized exclusively for victory, which frequently leads to strategies that diverge from human understanding and intuition. Consequently, a parallel research domain has emerged to develop models that replicate human decision-making patterns across various skill levels and individual styles (Levene and Fenner, 2011).

1.1 Maia: A First Step Toward Human-Like Chess Models

An initial methodology was established by McIlroy-Young et al. (2020) through the introduction of Maia. Maia adapts the AlphaZero convolutional residual network (Residual Network (ResNet)) backbone (He et al., 2016), replacing self-play reinforcement learning with supervised training on human game databases. Standard Monte Carlo Tree Search (MCTS) (Coulom, 2006) was omitted in the default configuration because unconstrained minimax lookahead reduced the accuracy of human move prediction. The network is trained on historical game records, taking the active board state and the preceding 12 plies as input to predict the subsequent move executed by the human player.

The Maia framework bins games according to the Elo rating of the participating players. Distinct models are trained on specific rating intervals, such as Maia 1100 for players rated between 1100 and 1199, Maia 1200 for 1200 to 1299, and continuing up to Maia 1900. Move prediction accuracy reaches approximately 50%, outperforming traditional chess engines such as Stockfish and LeelaChessZero when predicting human actions on those rating ranges, as summarized in Table 1.1.

1.2 Limitations of Maia

While Maia provides substantial predictive alignment, it presents notable structural limitations. At higher Elo ratings, model accuracy plateaus while engine evaluations become more predictive. Additionally, the network exhibits reduced accuracy when predicting human blunders, defined as moves decreasing the winning probability by 10% or more.

To address blunder modeling, the authors trained auxiliary networks explicitly tasked with classifying blunders using features such as rating and remaining clock time, achieving 71.7% classification accuracy. Furthermore, discrete rating bins do not accommodate strategic heterogeneity, as players sharing an identical rating often exhibit divergent stylistic preferences.

Table 1.1: Maximum move-matching accuracy for selected Maia, Leela, and Stockfish models (McIlroy-Young et al., 2020).

Model	**Best Test Set (Elo Rating)**	**Max Move Accuracy (%)**
Maia Models		
Maia 1100	1200	50.8%
Maia 1300	1400	51.8%
Maia 1500	1700	52.2%
Maia 1700	1800	52.7%
Maia 1900	1900	52.9%
Leela Models		
Leela 1000	1100	26.3%
Leela 1600	1100	31.4%
Leela 2200	1100	37.1%
Leela 2700	1400	40.2%
Leela 3200	1900	46.0%
Stockfish (SF)		
SF Depth 1	1900	39.1%
SF Depth 5	1900	35.4%
SF Depth 9	1900	39.8%
SF Depth 11	1900	40.6%
SF Depth 15	1900	41.1%

1.3 Maia-2: A Unified Model

To overcome the fragmentation of discrete models, Tang et al. (2024) introduced Maia-2. Maia-2 addresses the behavioral discontinuities observed between independent Maia models and incorporates opponent rating conditioning.

Maia-2 utilizes a unified architecture combining a ResNet convolutional backbone with a Vision Transformer (ViT) attention mechanism (Dosovitskiy et al., 2020). Output feature channels are processed as spatial patches, while the skill levels of both the active player and the opponent are encoded as continuous categorical embeddings injected into the query representations of a Skill-Aware Attention module. The processed representations feed three output heads: a move policy head, a position value head, and an auxiliary head predicting move properties such as checks and captures, as illustrated in Figure 1.1. While recent advances such as Chessformer (Monroe et al., 2026) achieve competitive move-matching accuracy by unifying chess tasks into dense transformer backbones, Maia-2 provides a modular, computationally lightweight convolutional-attention baseline specifically suited for localized parameter-efficient adapter intervention.

By modeling chess positions as a Markov Decision Process (MDP), Maia-2 processes individual board states rather than full 12-ply sequences, reducing computational requirements during training.

1.4 Results of Maia-2

Maia-2 achieves an average move prediction accuracy of 53.25%, outperforming Maia, LeelaChessZero, and Stockfish, as summarized in Table 1.2. Model perplexity, which quantifies predictive uncertainty by measuring the exponentiated cross-entropy over observed moves, decreases relative to Maia, as reported in Table 1.3.

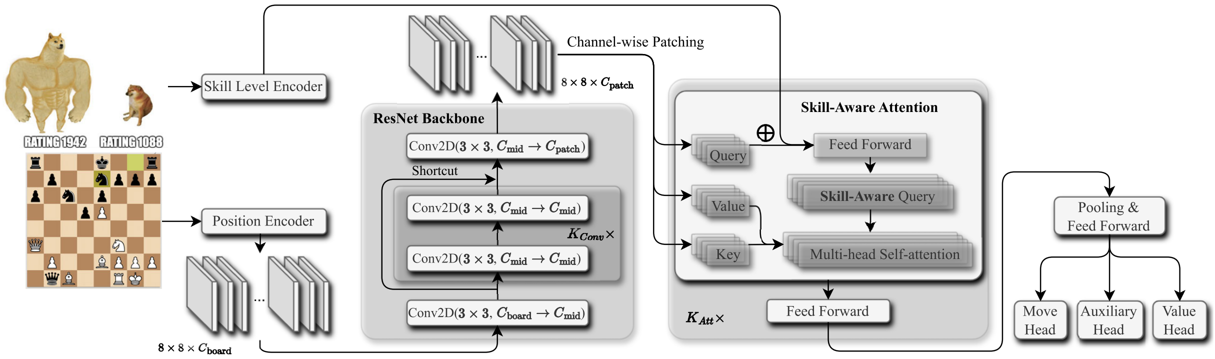


Figure 1.1: Architecture of the Maia-2 model (Tang et al., 2024).

Table 1.2: Move prediction accuracy of Maia-2 compared to Maia, Leela, and Stockfish (Tang et al., 2024).

Model	**Skilled (<1600)**	**Advanced (1600-2000)**	**Master (>2000)**	**Average (Avg)**
Stockfish 15	36.86%	39.83%	44.61%	40.43%
Leela 3200	39.97%	44.29%	47.75%	44.00%
Maia 1100	51.48%	49.13%	45.85%	48.82%
Maia 1500	50.79%	52.61%	50.76%	51.39%
Maia 1900	48.51%	52.26%	53.20%	51.32%
Maia-2	**51.72%**	**54.15%**	**53.87%**	**53.25%**

Table 1.3: Move prediction perplexity of Maia-2 compared to Maia (Tang et al., 2024).

Model	**Skilled (<1600)**	**Advanced (1600-2000)**	**Master (>2000)**	**Average (Avg)**
Maia 1100	5.05	5.56	6.01	5.54
Maia 1500	4.98	5.31	5.50	5.26
Maia 1900	4.44	4.71	4.86	4.67
Maia-2	**4.30**	**3.90**	**4.02**	**4.07**

Maia-2 demonstrates higher monotonicity and transitional consistency across rating progressions than Maia, as detailed in Table 1.4. A representative mate-in-1 puzzle evaluated across skill levels is illustrated in Figure 1.2.

Table 1.4: Percentage of monotone and transitional predictions for Maia-2 compared to Maia-1 (Tang et al., 2024).

Model	**% Monotone**			**% Transitional**		
	Skilled	**Advanced**	**Master**	**Skilled**	**Advanced**	**Master**
Maia-1	1.61%	1.42%	1.14%	13.34%	18.14%	20.48%
Maia-2	**27.61%**	**28.51%**	**26.38%**	**22.59%**	**23.39%**	**21.72%**

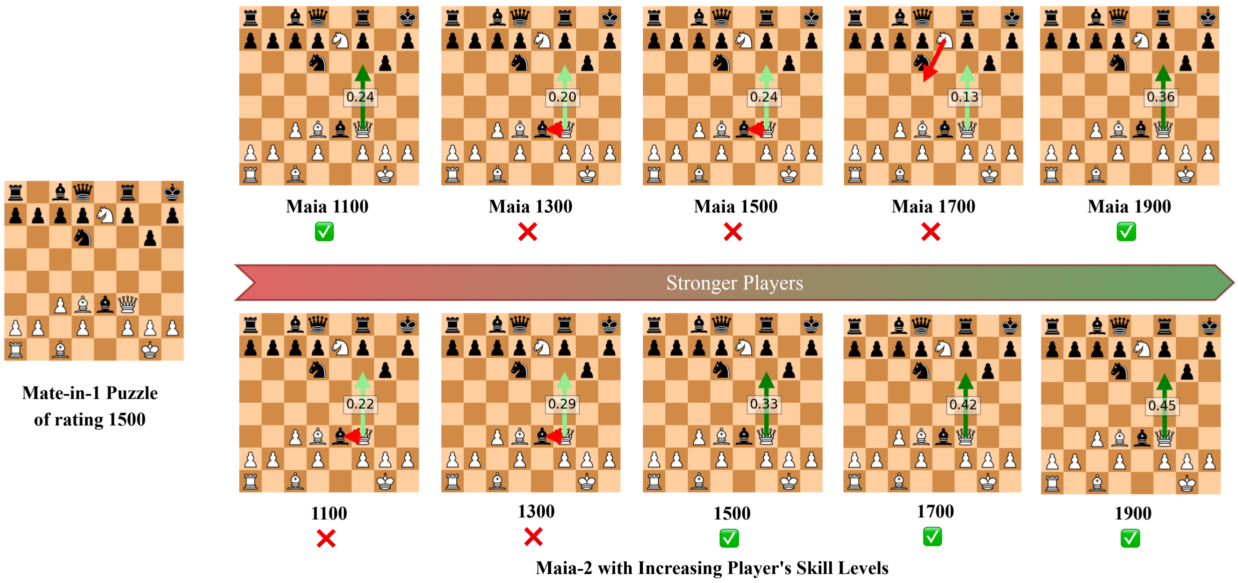


Figure 1.2: Example case of a 1500-rated Mate-in-1 puzzle where Maia-2 is coherent with the Elo rating, while Maia-1 is not (Tang et al., 2024).

The model adjusts move selections in response to active player and opponent ratings, as shown in Figure 1.3. Furthermore, linear probe analyses provide empirical evidence that intermediate layer activations encode tactical and positional features, as presented in Figure 1.4.

1.5 Maia4All: Modeling Individual Behavior

While Maia-2 captures population-level behavior across Elo ratings, individual players within identical rating brackets exhibit distinct tactical and strategic tendencies. Earlier individual adaptation frameworks, such as Maia Individual (McIlroy-Young et al., 2022), required approximately 5,000 games per player to avoid performance degradation.

To address low-resource individualization, Tang et al. (2026) introduced Maia4All. The framework establishes an enrichment stage using representative prototype players to bridge population embeddings with individual representations. During subsequent adaptation, a Prototype Matching Network (PMN) computes similarity weights between target player histories and prototype embeddings to initialize an individual prior. Fine-tuning is restricted to the individual embedding vector while backbone parameters remain frozen. Recent literature similarly highlights the efficacy of disentangled style vectors (Carlson, 2026) and generative style representations (Omi et al., 2025) to isolate individual strategic profiles.

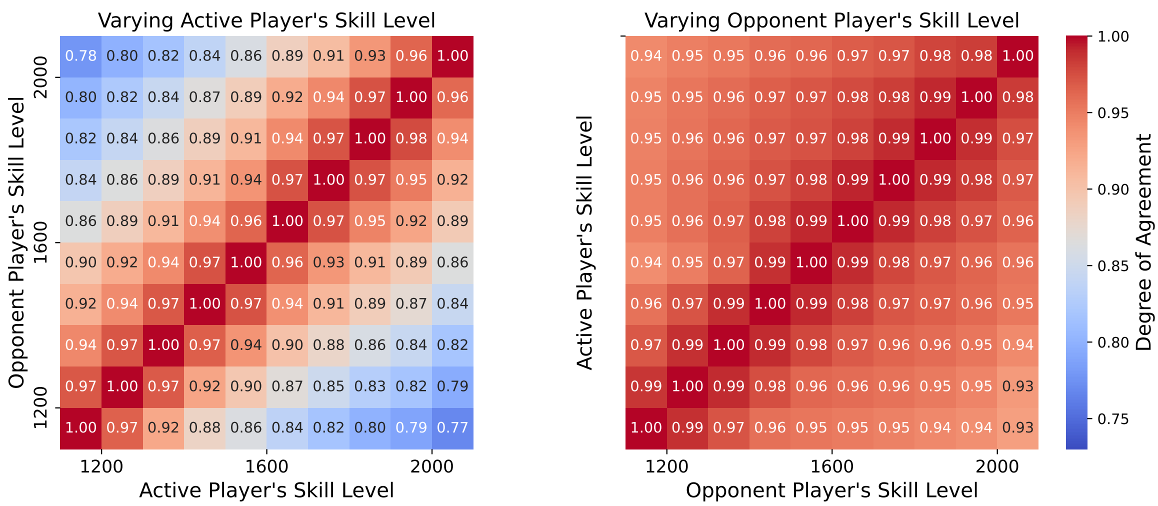


Figure 1.3: Agreement of Maia-2's move predictions when varying the active player's skill level (left) and the opponent's skill level (right) (Tang et al., 2024).

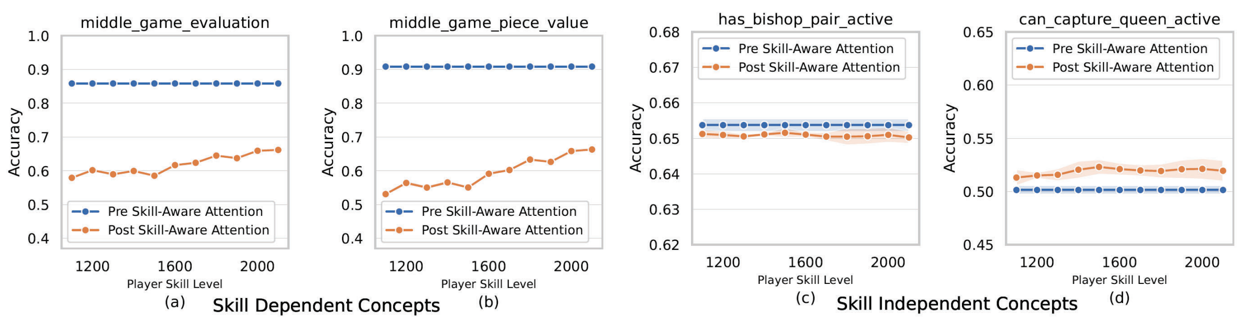


Figure 1.4: Probing results for level-dependent and level-independent concepts in Maia-2 (Tang et al., 2024).

1.6 Results of Maia4All

Table 1.5: Performance on unseen players under low-resource settings (Tang et al., 2026).

	Move Prediction Accuracy ↑				**Move Prediction Perplexity ↓**			
#Positions	20,000	8000	2000	800	20,000	8000	2000	800
#Games	≈500	≈200	≈50	≈20	≈500	≈200	≈50	≈20
Maia	0.5132	0.5132	0.5132	0.5132	5.4530	5.4530	5.4530	5.4530
Maia-2	0.5146	0.5146	0.5146	0.5146	4.5316	4.5316	4.5316	4.5316
Maia-2-Strength	0.5195	0.5196	0.5193	0.5189	4.4932	4.4939	4.4976	4.5022
Maia4All-Strength	0.5308	0.5298	0.5279	0.5249	4.3077	4.3238	4.3658	4.4151
Maia4All-Prototype	**0.5365**	**0.5348**	**0.5334**	**0.5322**	**4.2295**	**4.2431**	**4.2669**	**4.2988**

Table 1.6: Performance on unseen players with relatively rich histories (100,000 positions ≈ 2,500 games) (Tang et al., 2026).

	Move Prediction Accuracy ↑				**Move Prediction Perplexity ↓**			
Skill Categories	**Skilled**	**Advanced**	**Master**	**Overall**	**Skilled**	**Advanced**	**Master**	**Overall**
Maia	0.4996	0.5099	0.5285	0.5132	5.8687	5.4642	5.1300	5.4530
Maia-2	0.5008	0.5158	0.5364	0.5146	4.7900	4.4389	4.1936	4.5316
Maia-2-Strength	0.5071	0.5212	0.5400	0.5199	4.7264	4.4113	4.1760	4.4903
Maia4All-Strength	0.5226	0.5376	0.5478	0.5336	4.4504	4.1733	4.0291	4.2599
Maia4All-Prototype	**0.5261**	**0.5408**	**0.5554**	**0.5381**	**4.4018**	**4.1048**	**3.9219**	**4.1899**

As reported in Table 1.5, Maia4All achieves 53.22% move prediction accuracy with 20 recorded games (800 positions). Evaluated on players with larger archives ($\approx 2,500$ games), accuracy reaches 53.81%, as detailed in Table 1.6. Furthermore, the shared embedding representation enables stylometric player identification across candidates (McIlroy-Young et al., 2021).

1.7 Problem Statement and Objectives

Existing individualization methods have primarily been evaluated on amateur and intermediate players. Modeling world champions and elite competitors presents distinct challenges due to the tactical precision and depth of their decisions.

Additionally, standard evaluation methodologies rely heavily on top-1 move prediction accuracy. This metric enforces a deterministic assumption that fails to account for valid alternative moves and strategic variance. While probabilistic networks produce complete distribution vectors over candidate actions, top-1 evaluation collapses these distributions.

This research addresses two core objectives. The first objective focuses on extending parameter-efficient individual behavior modeling to a cohort of 14 elite chess players active in the 20th century. The second objective establishes evaluation methodologies that quantify behavioral alignment through discrete state-level divergence on controlled board positions (JSD_{action}) alongside a critical diagnostic of continuous latent manifold representations (JSD_{spatial}).

2. Methodology

The objective of this research is to develop and evaluate agents capable of reproducing player-specific decision behaviors for elite chess players. The methodology extends the pre-trained Maia-2 architecture across a target cohort of 14 elite players active in the 20th century, each possessing historical archives of at least 2,000 recorded games.

2.1 Benchmark Dataset and Cohort Characterization

Historical game records were compiled from the public archives of `Chessgames.com` (accessed February 2026). The candidate pool was restricted to players with substantial competitive volume across classical over-the-board tournament play. Positional records were verified using the `python-chess` library to eliminate corrupted notations and incomplete games. While an initial pool of 16 candidates was considered, two players were excluded following data sanitization due to incomplete move sequences and inconsistent tournament headers, yielding a finalized cohort of 14 elite players. Among them, six held the title of World Chess Champion (Alexander Alekhine, Mikhail Tal, Tigran Petrosian, Anatoly Karpov, Garry Kasparov, and Viswanathan Anand), while the remaining eight represent world-class players and Candidates tournament finalists (Alexander Beliavsky, Bent Larsen, Jan Timman, Lajos Portisch, Nigel Short, Ulf Andersson, Vassily Ivanchuk, and Viktor Korchnoi).

Table 2.1: Dataset characterization across the cohort of 14 elite players.

Player	Career Span	Total Games	White (%)	Black (%)	Total Positions
A. Alekhine	1905–1946	2,297	67.8%	32.2%	87,700
A. Beliavsky	1967–2020	2,834	50.6%	49.4%	108,668
A. Karpov	1961–2022	4,500	54.0%	46.0%	153,999
B. Larsen	1948–2008	2,896	50.4%	49.6%	115,589
G. Kasparov	1973–2025	3,098	55.2%	44.8%	96,202
J. Timman	1964–2019	4,166	49.9%	50.1%	143,462
L. Portisch	1952–2014	3,447	50.4%	49.6%	117,977
M. Tal	1949–1992	3,231	52.5%	47.5%	103,679
N. Short	1974–2025	3,256	51.7%	48.3%	118,327
T. Petrosian	1942–1983	2,211	51.5%	48.5%	68,403
U. Andersson	1967–2023	3,062	50.1%	49.9%	100,619
V. Ivanchuk	1983–2025	4,349	50.1%	49.9%	163,249
V. Korchnoi	1945–2015	5,332	51.1%	48.9%	195,682
V. Anand	1984–2026	4,528	51.0%	49.0%	165,823
Total	—	49,207	52.2%	47.8%	1,739,379

2.2 Behavior Generation Pipeline

2.2.1 Player-Specific Fine-Tuning

To individualize move predictions, the pre-trained Maia-2 model is adapted by instantiating a learnable player embedding for each elite player while keeping all backbone convolutional, attention, and head weights frozen, following general policy-value transfer principles (Soemers et al., 2023a). Embeddings are initialized using the terminal > 2000 Elo rating vector from Maia-2.

Hyperparameters for embedding fine-tuning are summarized in Table 2.2. The AdamW optimizer is employed with an initial learning rate of 1×10^{-2} decayed to 1×10^{-4} via a Cosine Annealing schedule over 5 epochs. Gradient clipping with a maximum norm of 1.0 is enforced. Cross-Entropy loss with a label smoothing factor of 0.05 is augmented with an L2 anchor penalty ($\lambda = 1 \times 10^{-5}$) regularizing the distance to the base > 2000 Elo embedding.

Hyperparameter	**Value**
Optimizer	AdamW
Learning Rate	1×10^{-2}
Weight Decay	1×10^{-4}
LR Scheduler	Cosine Annealing ($10^{-2} \rightarrow 10^{-4}$)
Batch Size	64
Epochs	5
Label Smoothing	0.05
L2 Anchor Weight	1×10^{-5}
Gradient Clipping (Max Norm)	1.0

Table 2.2: Hyperparameters utilized for the player-specific fine-tuning process.

To ensure sound evaluation, a 5-fold cross-validation scheme partitioned at the game level is systematically applied.

2.2.2 Mixture of Experts and Low-Rank Adaptation

As an alternative to input-level conditioning, a Mixture of Experts with Low-Rank Adaptation (MoE-LoRA) framework was designed to inject player-conditioned residual updates directly into the network logit space, building upon parameter-efficient adaptation and sparse routing architectures (Hu et al., 2021; Shazeer et al., 2017; Frisoni et al., 2026).

The module, `PlayerMoEAdapter`, uses $N = 8$ parallel LoRA expert blocks E_i, as illustrated in Figure 2.1. Each expert consists of a down-projection matrix $\mathbf{W}^{(i)}_{\text{down}} \in \mathbb{R}^{r \times d_v}$ ($r = 8$) initialized with a normal distribution ($\sigma = 0.02$) and an up-projection matrix $\mathbf{W}^{(i)}_{\text{up}} \in \mathbb{R}^{d_{\text{out}} \times r}$ initialized to zero:

$$E_i(\mathbf{v}) = \left(\mathbf{W}^{(i)}_{\text{up}} \mathbf{W}^{(i)}_{\text{down}} \mathbf{v}\right) \cdot \frac{\alpha}{\max(1, r)} \tag{2.1}$$

with scaling parameter $\alpha = 1.0$.

Each player p is assigned a learnable routing embedding $\mathbf{e}_p \in \mathbb{R}^{32}$. An intermediate hidden vector $\mathbf{v}$ extracted via a forward hook on Maia-2's `last_ln` layer is concatenated with $\mathbf{e}_p$ to form $[\mathbf{v}; \mathbf{e}_p]$. A two-layer gating network with a 128-dimensional hidden ReLU layer outputs softmax routing weights $\mathbf{g} \in \mathbb{R}^N$:

$$\mathbf{g} = \text{softmax}\left(\mathbf{W}_2 \cdot \text{ReLU}\left(\mathbf{W}_1 [\mathbf{v}; \mathbf{e}_p] + \mathbf{b}_1\right) + \mathbf{b}_2\right) \tag{2.2}$$

The residual update $\Delta\mathbf{z} = \sum_{i=1}^{N} g_i E_i(\mathbf{v})$ modifies the base logits:

$$\mathbf{z}_{\text{final}} = \mathbf{z}_{\text{base}} + \Delta\mathbf{z} \tag{2.3}$$

Training hyperparameters are detailed in Table 2.3. The model is optimized over 20 epochs with a weight decay of 1×10^{-2} and Cosine Annealing learning rate schedule decaying from 1×10^{-2} to 1×10^{-5}.

Hyperparameter	**Value**
Optimizer	AdamW
Learning Rate	1×10^{-2}
Weight Decay	1×10^{-2}
LR Scheduler	Cosine Annealing ($10^{-2} \to 10^{-5}$)
Batch Size	64
Epochs	20
Label Smoothing	0.05
Gradient Clipping (Max Norm)	1.0

Table 2.3: Hyperparameters utilized for the MoE-LoRA training process.

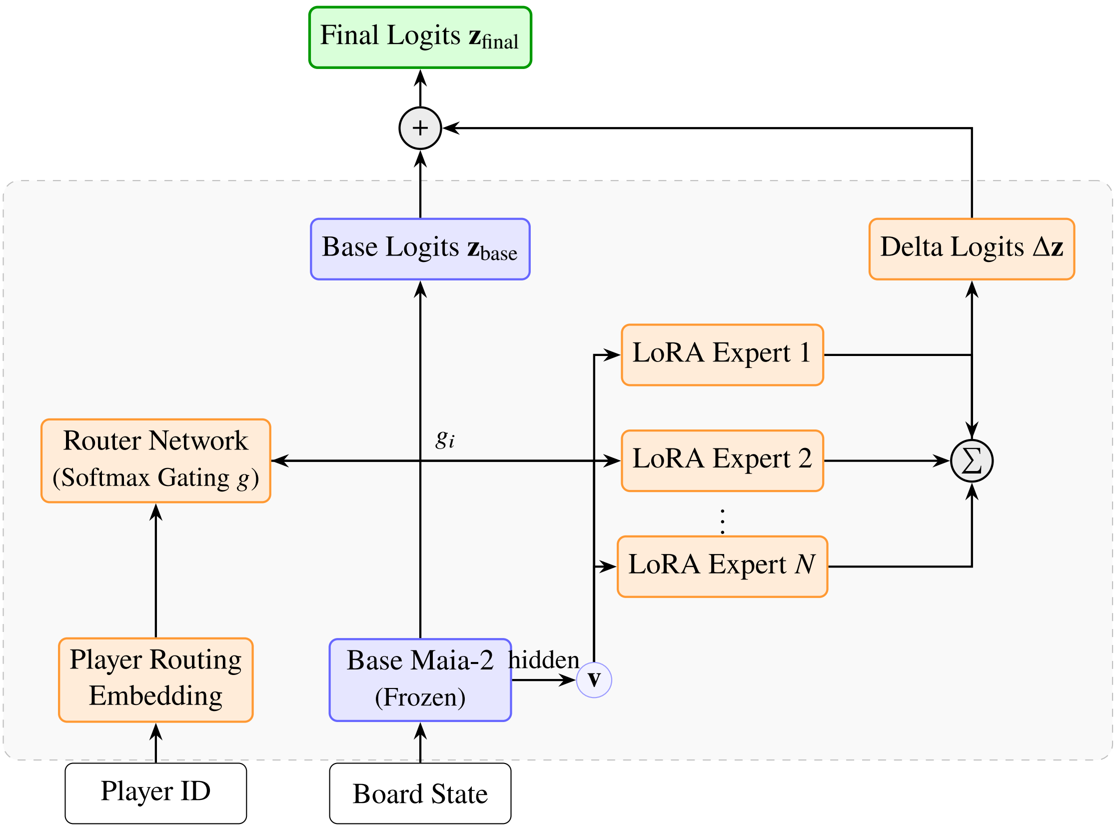


Figure 2.1: Architecture of the MoE-LoRA adapter module.

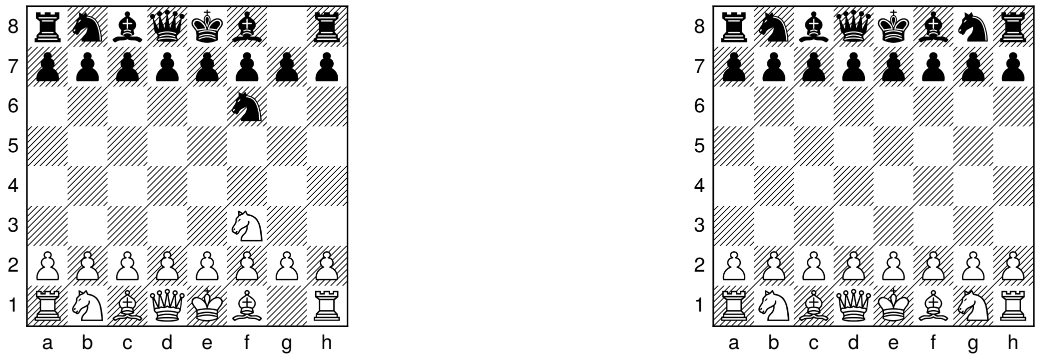
2.2.3 *Monte Carlo Tree Search Integration*

Direct neural policy generation occasionally exhibits tactical loop pathologies, including repetitive piece retreats leading to early draws by repetition (Figure 2.2) or tactical oversights (Figure 2.3). To address these vulnerabilities while maintaining human behavioral alignment, a lightweight

MCTS search using the PUCT selection algorithm is integrated, informed by search regularization and feature-guided lookahead principles (Soemers et al., 2019a,b; Jacob et al., 2022; Zhang et al., 2025; Rautureau et al., 2026a). Candidate action selection maximizes $Q(s, a) + U(s, a)$, where:

$$U(s, a) = c_{\text{puct}} \cdot P(s, a) \cdot \frac{\sqrt{N(s)}}{1 + N(a)} \tag{2.4}$$

with exploration constant $c_{\text{puct}} = 1.5$ and a shallow simulation budget strictly capped at $N = 50$ visits per decision.



(a) Position after 1. Nf3 Nf6 (b) Position after 2. Ng1 Ng8

Figure 2.2: Cyclic repetition pattern generated by standalone Maia-2.

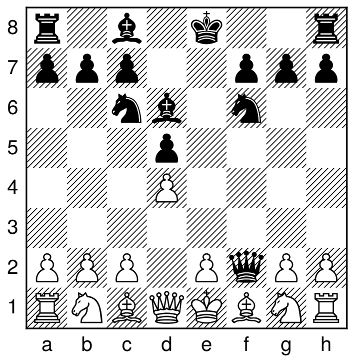


Figure 2.3: Opening blunder generated by unconstrained Maia-2 resulting in checkmate on move 8 (8... Qxf2#).

2.2.4 Descent Search Integration

As a deterministic search alternative, a value-guided Descent Search algorithm was implemented (Cohen-Solal and Cazenave, 2020). Transposition nodes maintain scalar evaluations v, game outcomes c, resolution flags r, and action counts $n(a)$. Child states are initialized using prior policy probabilities, and evaluations are updated via backward induction. The final policy $\pi(a \mid s)$ is derived through a softmax over child values scaled by temperature τ:

$$\pi(a \mid s) = \frac{\exp(v_a/\tau)}{\sum_b \exp(v_b/\tau)} \tag{2.5}$$

with $\tau = 1.0$ and an iteration budget capped at $N = 50$ search steps.

2.2.5 Nucleus Pruning Integration

To restrict the action space during search expansions, Nucleus Pruning based on top-p probability truncation (Holtzman et al., 2019) is applied. For legal moves sorted descending by prior probability, the minimal subset K satisfying cumulative mass threshold p is retained:

$$\sum_{i=1}^{K} P(s, a_i) \geq p \tag{2.6}$$

with $p = 0.95$.

2.3 Style Evaluation Pipeline

2.3.1 Standard Baselines: Move Accuracy and CP Error

Deterministic decision fidelity is quantified via Move Accuracy, measuring the percentage of exact top-1 action matches against historical moves. Tactical deviation is measured using Centipawn Error against the Stockfish 16 oracle (V_{best}), drawing upon tactical pattern extraction methods (Soemers et al., 2023c). Centipawn Loss is computed as $\text{CPL}(a) = \max(0, V_{\text{best}} - V(a))$, and Centipawn Error is defined as:

$$\text{CP Error} = \left| \frac{1}{|D|} \sum_{s \in D} \text{CPL}(a_{\text{player}}) - \frac{1}{|D|} \sum_{s \in D} \text{CPL}(a_{\text{model}}) \right| \tag{2.7}$$

2.3.2 Action-Level Jensen-Shannon Divergence on Common FENs (JSD_{action})

To evaluate policy calibration beyond deterministic move accuracy, the discrete **Action-Level Jensen-Shannon Divergence** (JSD_{action}) is calculated across legal move distributions (Lin, 1991). Given two discrete probability distributions P and Q defined over a shared legal action space $\mathcal{A}(s)$, the asymmetric **Kullback-Leibler (KL) divergence** is defined as:

$$D_{\text{KL}}(P \parallel Q) = \sum_{a \in \mathcal{A}(s)} P(a) \log_2 \left(\frac{P(a)}{Q(a)} \right) \tag{2.8}$$

Standard KL divergence presents two notable mathematical drawbacks in behavioral evaluation: it is asymmetric ($D_{\text{KL}}(P \parallel Q) \neq D_{\text{KL}}(Q \parallel P)$) and becomes unbounded ($D_{\text{KL}} \to +\infty$) whenever $Q(a) = 0$ for any action where $P(a) > 0$. To overcome these constraints, the symmetric Action-Level Jensen-Shannon Divergence is computed against the empirical mixture distribution $M = \frac{1}{2}(P_{\text{data}} + P_{\text{model}})$:

$$JSD_{\text{action}} \left(P_{\text{data}} \parallel P_{\text{model}}\right) = \frac{1}{2} D_{\text{KL}} \left(P_{\text{data}} \parallel M\right) + \frac{1}{2} D_{\text{KL}} \left(P_{\text{model}} \parallel M\right) \tag{2.9}$$

The Jensen-Shannon Divergence satisfies essential mathematical properties for behavioral stylometry. It provides complete symmetry ($JSD(P \parallel Q) = JSD(Q \parallel P)$), weighting observed player choices and model predictions identically. When evaluated with base-2 logarithms, the metric is strictly bounded in the range $[0, 1]$, where zero indicates identical distributions and unity corresponds to mutually disjoint supports. Because the mixture distribution satisfies $M(a) \geq \frac{1}{2}P(a) > 0$ across active actions, the logarithmic ratio remains bounded, ensuring numerical regularity without infinite values. Finally, the square root of the divergence, $d_{\text{JS}}(P, Q) = \sqrt{JSD(P \parallel Q)}$, obeys the triangle inequality, establishing a mathematically rigorous Jensen-Shannon distance metric.

Common FENs Benchmark To prevent Dirac delta collapse on single-occurrence states and enable consistent cross-player divergence heatmaps, JSD_{action} is strictly evaluated on the curated *Common FENs* benchmark. Formally, the set of common board positions $\mathcal{S}_{\text{common}}$ corresponds to the exact intersection of board configurations encountered across all players in the cohort:

$$\mathcal{S}_{\text{common}} = \bigcap_{p \in \mathcal{P}} \mathcal{S}_p \tag{2.10}$$

where $\mathcal{P}$ represents the cohort of 14 elite players and $\mathcal{S}_p$ is the set of unique board positions recorded for player p.

In this dataset, this shared intersection comprises exactly $|\mathcal{S}_{\text{common}}| = 154$ unique board states. Evaluating divergence over these shared positions is necessary to compute pairwise cross-player divergence heatmaps on a controlled basis. Furthermore, because every player in the cohort has reached these configurations, these positions constitute states where empirical move distributions are the most representative of real player behavior.

2.3.3 Spatial Manifold Divergence Pipeline ($JSD_{spatial}$)

To evaluate continuous transition distributions across entire game records, a spatial manifold evaluation pipeline was implemented, as illustrated in Figure 2.4. Board transitions ($s_t \to s_{t+1}$) are represented as 2304-dimensional vectors by concatenating pre- and post-move $8 \times 8 \times 18$ spatial tensors. A symmetric AutoEncoder ($2304 \to 1024 \to 512 \to 256 \to 128 \to \cdots \to 2304$) (Hinton and Salakhutdinov, 2006) is trained exclusively on reference human training transitions using the Adam optimizer with a learning rate of 10^{-3}, batch size of 2048, and seed 42 over 10 epochs. GPU-accelerated cuML UMAP ($n_{\text{neighbors}} = 15$, $d_{\text{min}} = 0.1$, seed 42) projects the 128-D latent embeddings onto a 2D plane (McInnes et al., 2018). The 2D coordinates are mapped onto a fixed 15×15 spatial grid (225 bins) bounded by global domain limits ($x_{\text{min}}, x_{\text{max}}, y_{\text{min}}, y_{\text{max}}$) derived from reference human games to ensure comparable probability support. Spatial divergence across the resulting 2D grid probabilities is evaluated via JSD_{spatial} and its metric distance $d_{\text{JS}} = \sqrt{JSD_{\text{spatial}}}$.

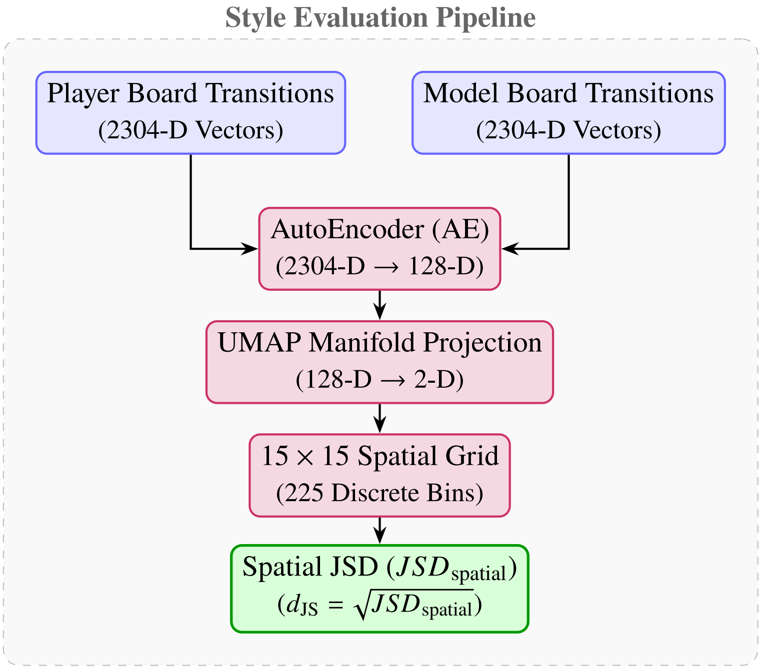


Figure 2.4: Overview of the Style Evaluation Pipeline.

3. Results and Discussion

All models and evaluation pipelines were implemented using the open-source codebase developed for this research[1]. Computations were executed on a workstation equipped with an AMD Ryzen 7 5700X eight-core processor, 32 GB of RAM, and an NVIDIA GeForce RTX 2060 GPU. A 5-fold cross-validation scheme partitioned at the game level is applied to all parameterized models. All ± intervals reported in summary tables denote the sample standard deviation across the 5 cross-validation folds.

3.1 Tactical Performance and Decision Accuracy

Table 3.1: Comprehensive performance summary across all model variants and evaluation dimensions.

Model Variant	Move Acc (%)	CP Error	JSD_{action} (Common)	$JSD_{spatial}$ (Common)
Direct Policy Variants				
Maia-2 Baseline	44.37%	56.01	0.4062	0.1512
Maia-2 FT	44.10%	52.67	0.3807	0.1543
Maia-2 Nucleus	44.37%	56.01	0.3988	0.1512
Maia-2 MoE-LoRA	45.87%	59.49	0.1996	**0.0409**
Descent Search Variants				
Maia-2 Descent	44.37%	56.12	0.7741	0.1512
Maia-2 N. + Descent	44.37%	56.05	0.5019	0.1512
Maia-2 FT + N. + Descent	43.41%	55.26	0.5389	0.1454
Maia-2 MoE-LoRA N. + Descent	45.91%	59.38	0.3512	0.0438
MCTS Search Variants				
Maia-2 MCTS	47.76%	25.39	0.3829	0.1500
Maia-2 N. + MCTS	47.71%	24.11	0.3834	0.1498
Maia-2 FT + N. + MCTS	46.77%	**22.84**	0.3544	0.1399
Maia-2 MoE-LoRA N. + MCTS	**49.09%**	28.22	**0.1861**	0.0434

To evaluate tactical strength and decision fidelity, Move Accuracy and Centipawn Error are evaluated. Table 3.1 summarizes overall performance, while per-player breakdowns are provided in Appendix A.7, A.8, A.9 for accuracy, and Appendix A.10, A.11, A.12 for Centipawn error.

3.1.1 Direct Policy Adaptation: Embedding Fine-Tuning vs. MoE-LoRA

In the direct policy setting without search, Maia-2 Baseline achieves a mean move accuracy of 44.37% ± 0.83% and a Centipawn error of 56.01 ± 7.63. Embedding fine-tuning alone (Maia-2 FT) yields 44.10% ± 0.91% accuracy and 52.67 ± 6.93 Centipawn error. Modifying a continuous conditioning vector within a frozen network provides limited capacity to displace shared decision boundaries toward individual styles.

[1] `https://github.com/SogliuzzoL/ChessBehaviors`

In contrast, Maia-2 MoE-LoRA increases direct policy accuracy to 45.87% ± 1.19%. Dynamic routing through rank-8 projection experts applies targeted adjustments to the logit space, improving move matching for players with distinct tactical profiles, such as Mikhail Tal (47.39%) and Anatoly Karpov (46.80%). The slight increase in Centipawn error to 59.49 ± 7.63 reflects the retention of human-aligned stylistic continuations that deviate from engine-optimal play.

3.1.2 Tactical Coherence and Search Integration

Integrating lightweight MCTS with a budget of $N = 50$ simulations eliminates cyclic piece shuttling and opening blunders, reducing Centipawn error from 56.01 to 24.11 ± 2.43 for Maia-2 N. + MCTS, and to 28.22 ± 3.42 for Maia-2 MoE-LoRA N. + MCTS, consistent with findings on regularized search (Soemers et al., 2020). Move accuracy reaches 49.09% ± 1.32%, exceeding 50% for Viswanathan Anand (51.10%) and Mikhail Tal (50.81%).

Conversely, shallow Descent Search shows limited effect on deterministic move accuracy (44.37% baseline, 45.91% MoE-LoRA), with Centipawn error remaining between 56.05 and 59.38. Because Descent Search derives policy probabilities via a softmax over propagated scalar values initialized from the prior, the top candidate predominantly matches the direct policy's highest-probability action.

3.1.3 Asymmetric Impact of Nucleus Pruning

Nucleus Pruning ($p = 0.95$) demonstrates asymmetric utility across search implementations. In Descent Search, pruning the low-probability tail truncates the effective branching factor, substantially accelerating search execution. In MCTS, pruning has negligible effect on runtime because the PUCT formula concentrates the 50-visit budget onto high-probability candidate branches, rarely expanding actions in the extreme probability tail.

3.2 Distributional Move Modeling and the Common FENs Paradigm

While top-1 move accuracy provides a deterministic metric, elite human players exhibit decision stochasticity across comparable positions. To evaluate full policy distribution matching, the symmetric Action-Level Jensen-Shannon Divergence (JSD_{action}) is evaluated.

3.2.1 Methodological Rationale: Support Sparsity and Shared Evaluation

Evaluating probability distributions across full career histories presents two primary limitations. First, because most board states appear only once in a player's record, empirical targets collapse into single-action Dirac vectors, penalizing probabilistic models that allocate valid mass across multiple sound moves. Second, assessing whether a model captures individual style requires evaluating all models across identical board states, preventing position-specific complexity from confounding stylistic comparison.

Consequently, distributional evaluations and cross-player divergence heatmaps are conducted on the benchmark of 154 Common FENs. Because these 154 board configurations are shared across all 14 historical elite players, they provide a dense, common basis over candidate actions, yielding representative empirical move distributions against which model policies can be evaluated reliably.

3.2.2 Quantitative Distributional Alignment on Common Positions

On Common FENs, Maia-2 Baseline achieves an average JSD_{action} of 0.4062 ± 0.0198, and Maia-2 FT records 0.3807 ± 0.0216. Maia-2 MoE-LoRA reduces mean divergence to 0.1996 ± 0.0290. This reduction is prominent for Anatoly Karpov (0.4132 to 0.1318), Jan Timman (0.3967 to 0.1638), and Alexander Beliavsky (0.4277 to 0.1782).

3.2.3 Impact of Search Policies on Probability Distributions

Unconstrained Descent Search distorts the softmax policy distribution through backward value propagation, increasing mean JSD_{action} to 0.7741 ± 0.0126. Applying Nucleus Pruning restores calibration to 0.5019 ± 0.0184 for Maia-2 N. + Descent, and 0.3512 ± 0.0342 when combined with MoE-LoRA. In MCTS, PUCT visit count distributions preserve probabilistic alignment (0.3834 ± 0.0200), and Maia-2 MoE-LoRA N. + MCTS achieves the lowest overall divergence at 0.1861 ± 0.0225.

3.2.4 Inter-Player Stylometric Heatmap Analysis

Cross-player divergence heatmaps are constructed by evaluating models against the empirical move distributions of all 14 historical elite players on Common FEN positions.

Figure 3.1 presents the empirical ground truth matrix measuring divergence directly between players. The diagonal exhibits an exact divergence of 0.000. Tactical players such as Vassily Ivanchuk and Garry Kasparov display relatively low mutual divergence (0.096), whereas contrasting styles, such as Garry Kasparov and Alexander Alekhine, exhibit higher divergence (0.205).

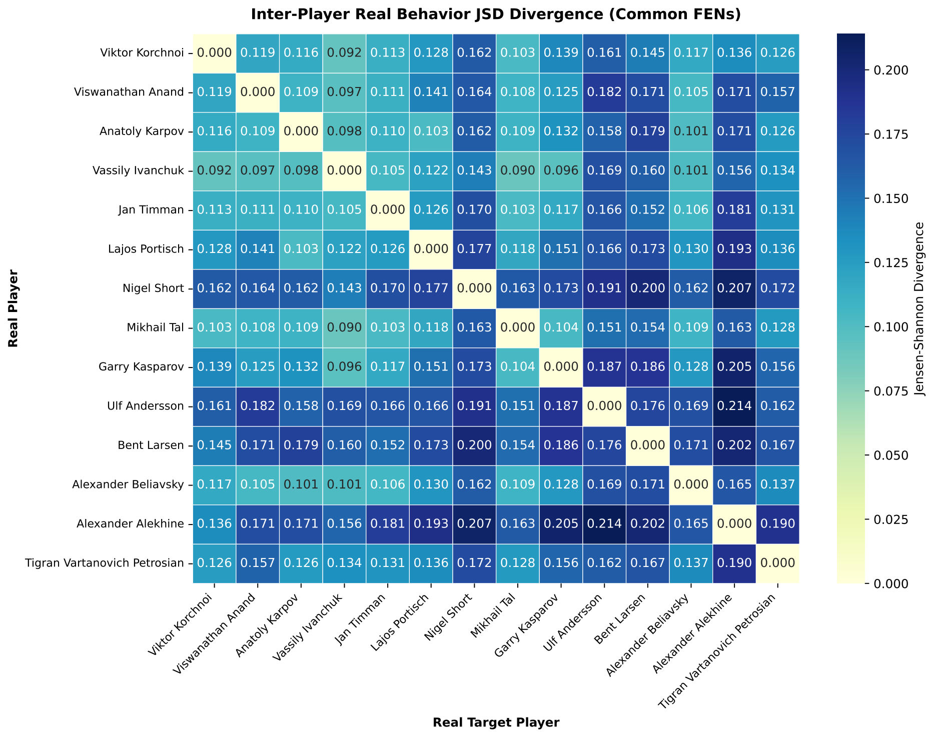


Figure 3.1: Empirical Inter-Player JSD_{action} matrix on Common FEN positions.

Figure 3.2 shows the alignment matrix for Maia-2 Baseline. Without player conditioning, every row displays identical horizontal values (0.261 for Viktor Korchnoi, 0.286 for Anatoly Karpov, 0.291 for Alexander Alekhine), confirming that an unconditioned model outputs a static average distribution.

When MoE-LoRA adapters are deployed, the cross-divergence matrices transition into structured matrices featuring a pronounced diagonal (Figure 3.3).

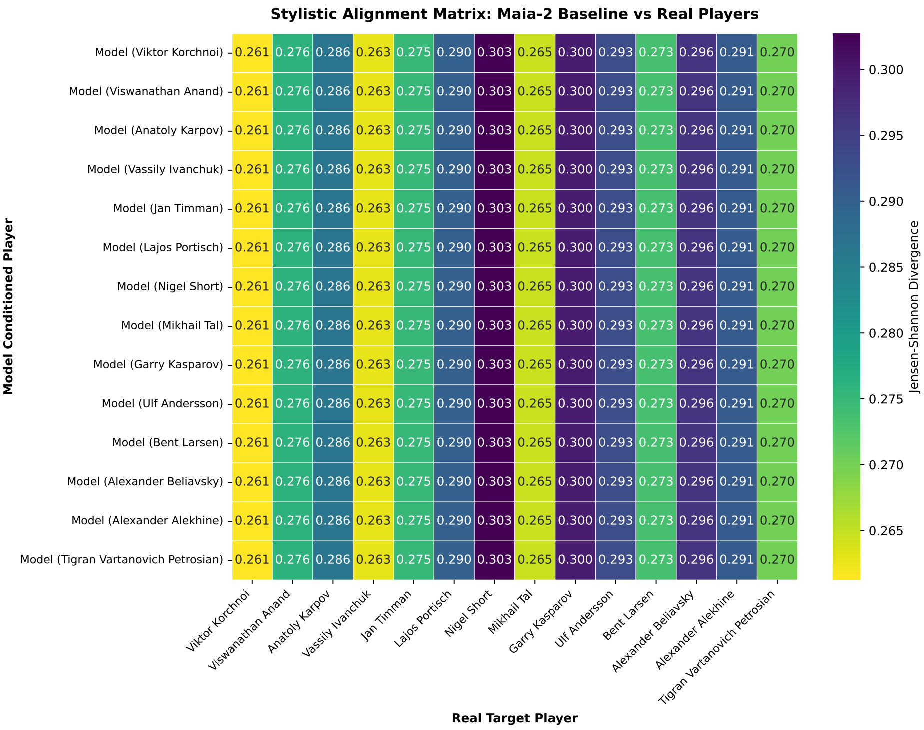


Figure 3.2: Stylistic alignment matrix for Maia-2 Baseline across real player move distributions on Common FENs, exhibiting uniform horizontal banding.

Diagonal entries are minimized for most players, including Anatoly Karpov (0.091 in MoE-LoRA, 0.109 in MCTS), Viktor Korchnoi (0.132 and 0.112), Mikhail Tal (0.132 and 0.106), and Ulf Andersson (0.130 and 0.123). Certain edge cases exhibit close cross-divergences. For Garry Kasparov, the model records 0.154 against Kasparov's ground truth, with proximate values on Vassily Ivanchuk (0.140) and Mikhail Tal (0.147), a pattern maintained under MCTS (0.137 on Kasparov, 0.127 on Ivanchuk, 0.130 on Tal). For Nigel Short, diagonal divergence (0.168 in MoE-LoRA, 0.163 in MCTS) is close to values against Korchnoi (0.164 and 0.158) and Tal (0.173 and 0.156).

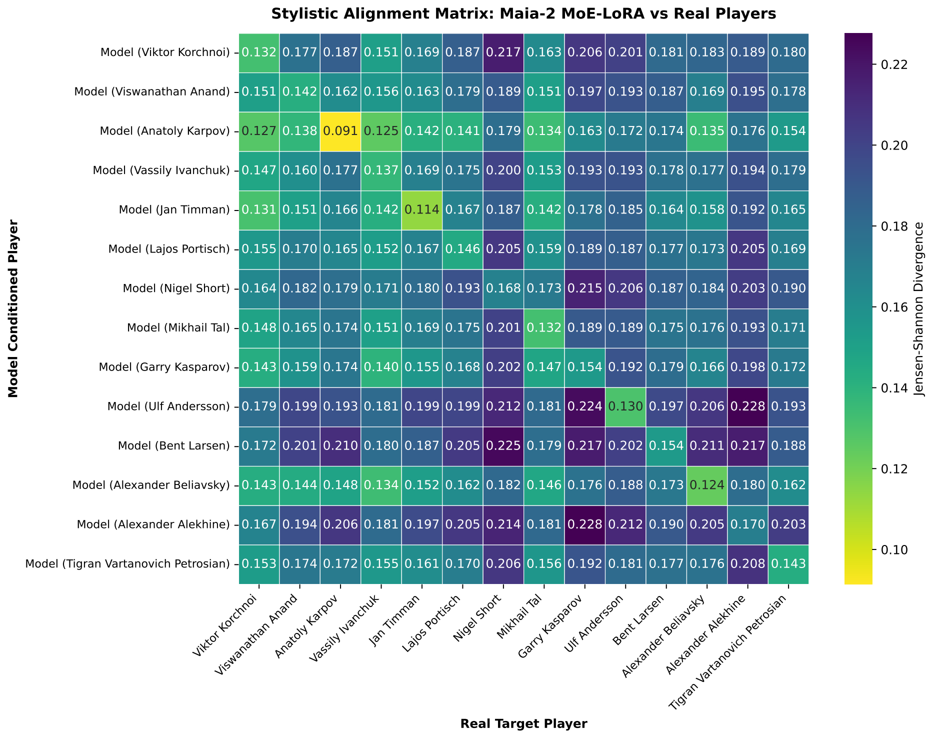
These cross-divergence overlaps mirror the empirical ground truth relationships in Figure 3.1, where Kasparov, Tal, and Ivanchuk exhibit low mutual divergence (0.090 to 0.104). Because their historical move distributions on these shared board positions are statistically close, model predictions reflect this proximity. Conversely, players with distinctive distribution profiles, such as Ulf Andersson, are isolated with clear contrast.

3.3 Stylistic Manifold Evaluation and Critical Diagnostic

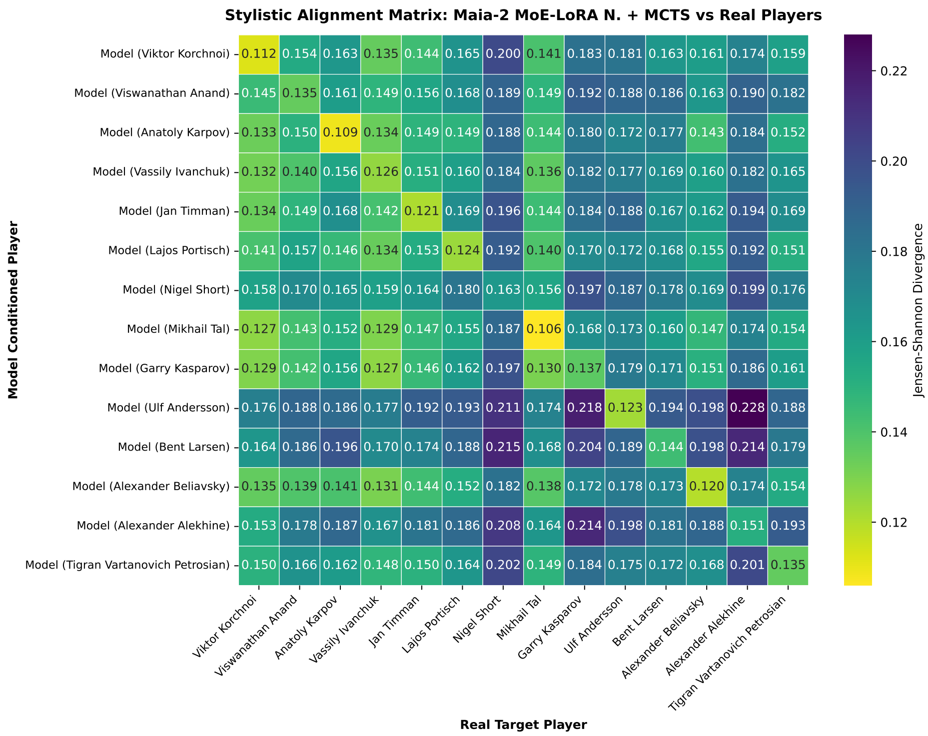
The stylistic evaluation pipeline was designed to project 2304-dimensional board transitions into a compressed 2D latent space via an AutoEncoder and UMAP, discretizing the manifold into a 15×15 grid to compare movement distributions across complete game histories using Spatial Manifold Divergence (JSD_{spatial}).

3.3.1 Apparent Discrimination on the Full Game Dataset

On the complete game dataset, the manifold metric yields low JSD_{spatial} across all variants, averaging 0.0069 ± 0.0017 under MoE-LoRA (Appendix A.6). As shown in Figure 3.4a, the alignment matrix for Maia-2 Baseline exhibits an apparent diagonal structure, recording values of 0.016 for Bent Larsen and 0.022 for Ulf Andersson, while off-diagonal values reach 0.333.

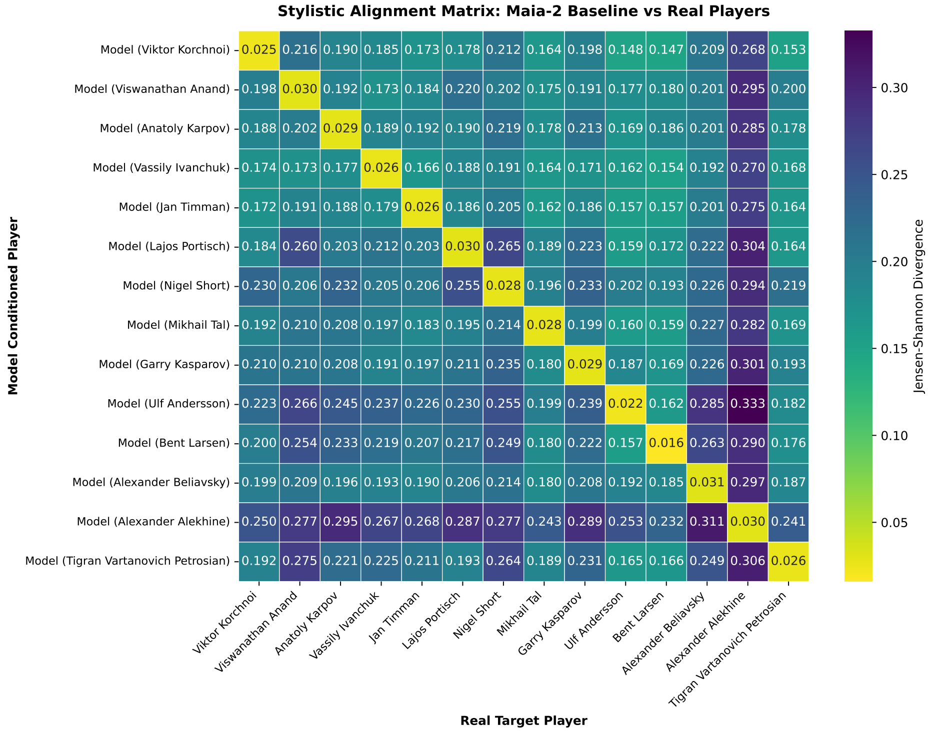


(a) Maia-2 MoE-LoRA (Direct Policy)

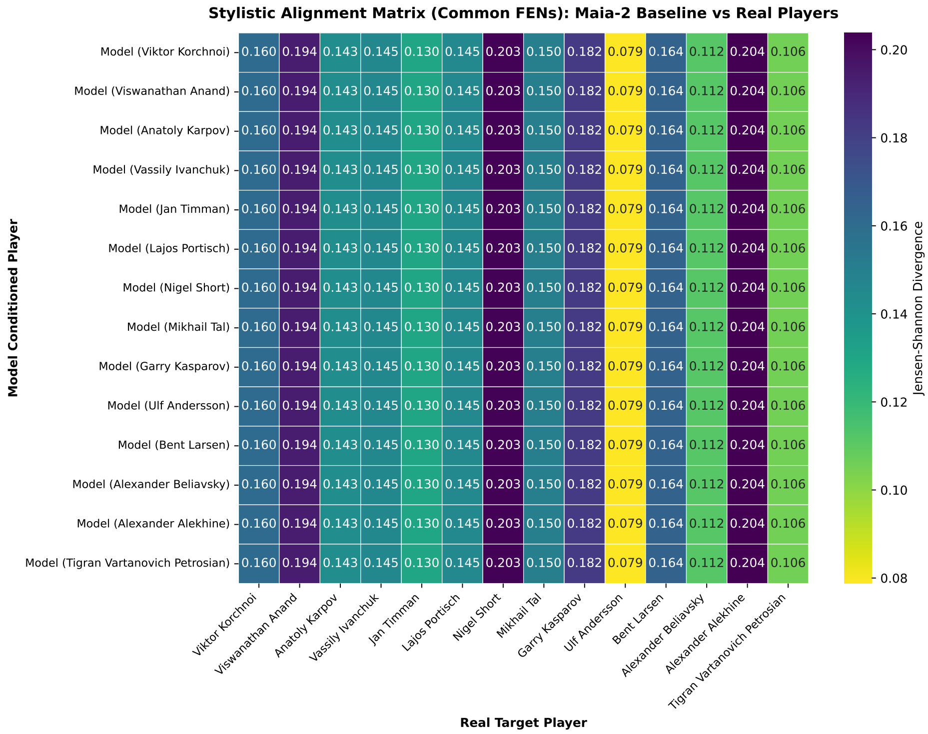


(b) Maia-2 MoE-LoRA N. + MCTS (Search Augmented)

Figure 3.3: Cross-player JSD_{action} alignment matrices on Common FEN positions comparing direct policy MoE-LoRA against the search-augmented MoE-LoRA MCTS model.



(a) Full Dataset Evaluation (Apparent Diagonal Artifact)



(b) Common FENs Evaluation (Collapse of Spatial Differentiation)

Figure 3.4: Diagnostic comparison of the Stylistic Alignment Matrix (JSD_{spatial}) for Maia-2 Baseline evaluated on the complete game dataset versus the controlled Common FENs benchmark.

3.3.2 Failure of Manifold Discrimination on Common Positions

When restricted to Common FEN positions, the diagonal structure vanishes entirely (Figure 3.4b), reverting to uniform horizontal rows identical across columns (0.160 for Viktor Korchnoi, 0.079 for Ulf Andersson, 0.204 for Alexander Alekhine). On identical positions, the manifold representations generated for different player profiles become indistinguishable.

3.3.3 Diagnostic of Positional Context Bias

This empirical divergence reveals that the AutoEncoder and UMAP pipeline primarily encodes macro-level board configurations, piece placement, and pawn structures rather than local move-selection policies. On full game histories, the metric captures differences in opening repertoires and reached pawn structures. Because different elite players favor distinct opening branches, their game datasets form disjoint clusters in the 2304-dimensional space, which the AutoEncoder separates, producing an apparent diagonal even with an unconditioned baseline model.

When positional distributions are held constant via Common FENs, spatial projection cannot isolate decision nuances. This demonstrates that stylistic evaluation in discrete domains requires measuring action probability distributions on controlled board states (JSD_{action}) rather than unconstrained state trajectories, underscoring the importance of localized game features (Soemers et al., 2023b; Piette et al., 2021).

4. Conclusion and Future Work

This research investigated the modeling and evaluation of player-specific decision behaviors in chess, extending human-aligned generative methods to elite 20th-century competitors and world champions. Building upon the unified architecture of Maia-2, the work explored parameter-efficient adaptation, search regularization, and distributional evaluation methodologies.

4.1 Summary of Contributions

The primary architectural contribution is the demonstration of Mixture of Experts with Low-Rank Adaptation (MoE-LoRA) for elite player behavioral alignment. While continuous embedding fine-tuning remains constrained by the frozen backbone, the MoE-LoRA framework dynamically routes intermediate activations through specialized rank-8 experts. This increases direct policy move prediction accuracy to 45.87% and reduces Action-Level JSD (JSD_{action}) on shared opening positions from 0.4062 to 0.1996.

Integrating lightweight MCTS ($N = 50$, $c_{\text{puct}} = 1.5$) regularized by Nucleus Pruning ($p = 0.95$) resolved cyclic repetition patterns and opening blunders, reducing Centipawn error from 56.01 to 28.22 while achieving a peak top-1 accuracy of 49.09%. Descent Search proved less effective for deterministic move matching, while Nucleus Pruning provided computational speedups during recursive expansions.

4.2 Methodological Insights and Limitations

This research highlights methodological considerations for behavioral evaluation in sequential domains. Top-1 move accuracy enforces a deterministic assumption that overlooks valid strategic alternatives, while full-corpus distributional evaluations suffer from empirical support sparsity. Evaluating discrete action probabilities on Common FENs (JSD_{action}) resolved these limitations, producing cross-player stylometric matrices that identify personalized decision tendencies.

Furthermore, this work identified a positional context bias in spatial manifold evaluation pipelines (JSD_{spatial}). Compressing full board transitions into a 2D latent space via AutoEncoders and UMAP reflected opening repertoire differences rather than move-selection style, causing apparent player separation on full game records to collapse on shared positions. Valid behavioral stylometry in discrete games requires evaluating action probability distributions conditioned on controlled board states.

4.3 Future Work

The experimental findings identify several direct technical extensions regarding data efficiency, architecture optimization, algorithmic search dynamics, longitudinal behavioral shifts, and empirical validation.

The candidate selection criterion required historical archives of at least 2,000 games to ensure adequate training depth. Extending this framework to low-resource regimes can be pursued by integrating the prototype matching mechanisms established in Maia4All. Initializing expert routing parameters through weighted projections over informative prototype embeddings would enable adaptation to modern masters, historical players with sparse archives, or online practitioners with restricted game histories.

While parameter-efficient adaptation was utilized to prevent catastrophic forgetting and preserve general chess representations, evaluating full-network fine-tuning without freezing the underlying Maia-2 backbone remains an open comparison. Evaluating unconstrained gradient updates across convolutional and attention layers against low-rank modular updates will determine whether full-network optimization captures higher-order strategic patterns or primarily accelerates overfitting to opening repertoires.

The search hyperparameters implemented in the MCTS pipeline, specifically the visit budget of $N = 50$ and exploration constant $c_{\text{puct}} = 1.5$, were selected empirically to enforce a minimal tactical safety bound. Conducting a systematic parameter sweep across varying simulation budgets and exploration constants is necessary to quantify the exact Pareto frontier between tactical error reduction and distributional deviation from human move selections, leveraging rigorous agent selection methods (Stephenson et al., 2026).

The present methodology aggregated player records across entire careers into single static profiles. Because competitive chess players undergo marked strategic shifts throughout their careers, segmenting historical corpora into chronological career phases, corresponding to early formative periods, peak championship reigns, and late career stages, will allow the training of temporally conditioned models to track longitudinal stylistic evolution.

Evaluating the empirical playing strength of the adapted agents through round-robin tournament simulations and calibrated Elo benchmarking will verify whether the generated policies reproduce historical competitive dynamics, connecting with human-like general game playing frameworks such as CogniPlay (Rautureau and Piette, 2025) and Ludii (Piette et al., 2020). In parallel, conducting blinded perceptual studies with titled human players against various rating levels will evaluate whether human practitioners subjectively identify player-specific characteristics in interactive play, drawing upon human-agent interaction evaluation paradigms (Rosemarin and Rosenfeld, 2019).

Bibliography

Murray Campbell, A Joseph Hoane Jr, and Feng-hsiung Hsu. Deep blue. *Artificial intelligence*, 134(1-2):57–83, 2002.

Jason Carlson. Matilda: Engine-agnostic search with human policy guidance, 2026.

Quentin Cohen-Solal and Tristan Cazenave. Minimax strikes back. *arXiv preprint arXiv:2012.10700*, 2020.

Rémi Coulom. Efficient selectivity and backup operators in monte-carlo tree search. In *International conference on computers and games*, pages 72–83. Springer, 2006.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

Giacomo Frisoni, Lorenzo Molfetta, Davide Freddi, and Gianluca Moro. Mixture of masters: Sparse chess language models with player routing. *arXiv preprint arXiv:2602.04447*, 2026.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*, 2019.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.

Athul Paul Jacob, David J Wu, Gabriele Farina, Adam Lerer, Hengyuan Hu, Anton Bakhtin, Jacob Andreas, and Noam Brown. Modeling strong and human-like gameplay with kl-regularized search. In *International Conference on Machine Learning*, pages 9695–9728. PMLR, 2022.

Mark Levene and Trevor Fenner. A methodology for learning players' styles from game records. *International Journal of Artificial Intelligence and Soft Computing*, 2(4):272–286, 2011.

Jianhua Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information theory*, 37(1):145–151, 1991.

Reid McIlroy-Young, Siddhartha Sen, Jon Kleinberg, and Ashton Anderson. Aligning superhuman ai with human behavior: Chess as a model system. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1677–1687, 2020.

Reid McIlroy-Young, Yu Wang, Siddhartha Sen, Jon Kleinberg, and Ashton Anderson. Detecting individual decision-making style: Exploring behavioral stylometry in chess. *Advances in Neural Information Processing Systems*, 34:24482–24497, 2021.

Reid McIlroy-Young, Russell Wang, Siddhartha Sen, Jon Kleinberg, and Ashton Anderson. Learning models of individual behavior in chess. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1253–1263, 2022.

Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

Daniel Monroe, George Eilender, Philip Chalmers, Zhenwei Tang, and Ashton Anderson. Chessformer: A unified architecture for chess modeling. In *The Fourteenth International Conference on Learning Representations*, 2026.

Nabil Omi, Lucas Caccia, Anurag Sarkar, Jordan T Ash, and Siddhartha Sen. Generative modeling of individual behavior at scale. *arXiv preprint arXiv:2502.14998*, 2025.

Eric Piette, Dennis JNJ Soemers, Matthew Stephenson, Chiara F Sironi, Mark HM Winands, and Cameron Browne. Ludii–the ludemic general game system. In *ECAI 2020: 24th European Conference on Artificial Intelligence, 29 August–8 September 2020, Santiago de Compostela, Spain–Including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020)*, pages 411–418. SAGE Publications 1 Oliver's Yard, 55 City Road, London, EC1Y 1SP, 2020.

Eric Piette, Matthew Stephenson, Dennis JNJ Soemers, and Cameron Browne. General board game concepts. In *2021 IEEE Conference on Games (CoG)*, pages 01–08. IEEE, 2021.

Éric Piette, Walter Crist, Dennis JNJ Soemers, Lisa Rougetet, Summer Courts, Tim Penn, and Achille Morenville. Gametable cost action: Kickoff report. *ICGA Journal*, 46(1):11–27, 2024.

Aloïs Rautureau and Éric Piette. Cogniplay: a work-in-progress human-like model for general game playing. *arXiv preprint arXiv:2507.05868*, 2025.

Aloïs Rautureau, Tristan Cazenave, and Éric Piette. Generalized rapid action value estimation in memory-constrained environments. *arXiv preprint arXiv:2602.23318*, 2026a.

Aloïs Rautureau, Éric Piette, Tim Penn, Summer Courts, James Goodman, Walter Crist, and Dennis JNJ Soemers. Game ai for cultural heritage: Outcomes from the gametable third grant period opening meeting. *ICGA Journal*, 48(1):29–36, 2026b.

Hanan Rosemarin and Ariel Rosenfeld. Playing chess at a human desired level and style. In *Proceedings of the 7th international conference on human-agent interaction*, pages 76–80, 2019.

Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. *arXiv preprint arXiv:1701.06538*, 2017.

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017.

Dennis JNJ Soemers, Éric Piette, and Cameron Browne. Biasing mcts with features for general games. In *2019 IEEE Congress on Evolutionary Computation (CEC)*, pages 450–457. IEEE, 2019a.

Dennis JNJ Soemers, Eric Piette, Matthew Stephenson, and Cameron Browne. Learning policies from self-play with policy gradients and mcts value estimates. In *2019 IEEE Conference on Games (CoG)*, pages 1–8. IEEE, 2019b.

Dennis JNJ Soemers, Eric Piette, Matthew Stephenson, and Cameron Browne. Manipulating the distributions of experience used for self-play learning in expert iteration. In *2020 IEEE Conference on Games (CoG)*, pages 245–252. IEEE, 2020.

Dennis JNJ Soemers, Vegard Mella, Éric Piette, Matthew Stephenson, Cameron Browne, and Olivier Teytaud. Towards a general transfer approach for policy-value networks. *Transactions on Machine Learning Research*, 2023a.

Dennis JNJ Soemers, Eric Piette, Matthew Stephenson, and Cameron Browne. Spatial state-action features for general games. *Artificial Intelligence*, 321:103937, 2023b.

Dennis JNJ Soemers, Spyridon Samothrakis, Éric Piette, and Matthew Stephenson. Extracting tactics learned from self-play in general games. *Information Sciences*, 624:277–298, 2023c.

Loris Sogliuzzo, Aloïs Rautureau, and Eric Piette. Toward modeling player-specific chess behaviors. *arXiv preprint arXiv:2605.11893*, 2026.

Matthew Stephenson, Alex Newcombe, Éric Piette, and Dennis JNJ Soemers. Best agent identification for general game playing. *IEEE Access*, 2026.

Zhenwei Tang, Difan Jiao, Reid McIlroy-Young, Jon Kleinberg, Siddhartha Sen, and Ashton Anderson. Maia-2: A unified model for human-ai alignment in chess. *Advances in Neural Information Processing Systems*, 37:20919–20944, 2024.

Zhenwei Tang, Difan Jiao, Eric Xue, Reid McIlroy-Young, Jon Kleinberg, Siddhartha Sen, and Ashton Anderson. Learning to imitate with less: Efficient individual behavior modeling in chess. *Transactions on Machine Learning Research*, 2026. ISSN 2835-8856.

Yiming Zhang, Athul Paul Jacob, Vivian Lai, Daniel Fried, and Daphne Ippolito. Human-aligned chess with a bit of search. In *The Thirteenth International Conference on Learning Representations*, 2025.

A. Detailed Experimental Results

This appendix provides numerical evaluations across all 12 model configurations and 14 historical elite players. The tables present aggregated global summaries and per-player breakdowns for direct policy variants, Descent Search configurations, and Monte Carlo Tree Search implementations.

A.1 Global Performance Summaries

Table A.1: Overall Move Accuracy summary across model variants.

Model Variant	Mean Accuracy (%)	Std Dev (%)
Maia-2 Baseline	44.37%	± 0.83%
Maia-2 FT	44.10%	± 0.91%
Maia-2 Nucleus	44.37%	± 0.83%
Maia-2 MoE-LoRA	45.87%	± 1.19%
Maia-2 Descent	44.37%	± 0.83%
Maia-2 N. + Descent	44.37%	± 0.83%
Maia-2 FT + N. + Descent	43.41%	± 0.83%
Maia-2 MoE-LoRA N. + Descent	45.91%	± 1.27%
Maia-2 MCTS	47.76%	± 0.97%
Maia-2 N. + MCTS	47.71%	± 0.97%
Maia-2 FT + N. + MCTS	46.77%	± 0.95%
Maia-2 MoE-LoRA N. + MCTS	**49.09%**	± **1.32%**

Table A.2: Overall Centipawn Error (CP Error) summary across model variants.

Model Variant	Mean CP Error	Std Dev
Maia-2 Baseline	56.01	± 7.63
Maia-2 FT	52.67	± 6.93
Maia-2 Nucleus	56.01	± 7.63
Maia-2 MoE-LoRA	59.49	± 7.63
Maia-2 Descent	56.12	± 7.64
Maia-2 N. + Descent	56.05	± 7.59
Maia-2 FT + N. + Descent	55.26	± 7.47
Maia-2 MoE-LoRA N. + Descent	59.38	± 8.32
Maia-2 MCTS	25.39	± 2.51
Maia-2 N. + MCTS	24.11	± 2.43
Maia-2 FT + N. + MCTS	**22.84**	± **2.04**
Maia-2 MoE-LoRA N. + MCTS	28.22	± 3.42

Table A.3: Overall Action-Level Jensen-Shannon Divergence (JSD_{action}) summary on common FEN positions across model variants.

Model Variant	**Mean JSD$_{\text{action}}$**	**Std Dev**
Maia-2 Baseline	0.4062	± 0.0198
Maia-2 FT	0.3807	± 0.0216
Maia-2 Nucleus	0.3988	± 0.0195
Maia-2 MoE-LoRA	0.1996	± 0.0290
Maia-2 Descent	0.7741	± 0.0126
Maia-2 N. + Descent	0.5019	± 0.0184
Maia-2 FT + N. + Descent	0.5389	± 0.0193
Maia-2 MoE-LoRA N. + Descent	0.3512	± 0.0342
Maia-2 MCTS	0.3829	± 0.0200
Maia-2 N. + MCTS	0.3834	± 0.0200
Maia-2 FT + N. + MCTS	0.3544	± 0.0242
Maia-2 MoE-LoRA N. + MCTS	**0.1861**	± **0.0225**

Table A.4: Overall Action-Level Jensen-Shannon Divergence (JSD_{action}) on full game records across model variants.

Model Variant	**Mean JSD$_{\text{action}}$**	**Std Dev**
Maia-2 Baseline	0.5198	± 0.0075
Maia-2 FT	0.5234	± 0.0082
Maia-2 Nucleus	0.5122	± 0.0076
Maia-2 MoE-LoRA	0.5365	± 0.0072
Maia-2 Descent	0.8424	± 0.0035
Maia-2 N. + Descent	0.5940	± 0.0067
Maia-2 FT + N. + Descent	0.7027	± 0.0066
Maia-2 MoE-LoRA N. + Descent	0.6033	± 0.0069
Maia-2 MCTS	0.4792	± 0.0087
Maia-2 N. + MCTS	**0.4787**	± **0.0088**
Maia-2 FT + N. + MCTS	0.5166	± 0.0090
Maia-2 MoE-LoRA N. + MCTS	0.4963	± 0.0090

Table A.5: Overall Spatial Manifold Divergence (JSD_{spatial}) on common FEN positions across model variants.

Model Variant	**Mean JSD$_{\text{spatial}}$**	**Std Dev**
Maia-2 Baseline	0.1512	± 0.0358
Maia-2 FT	0.1543	± 0.0393
Maia-2 Nucleus	0.1512	± 0.0358
Maia-2 MoE-LoRA	**0.0409**	± **0.0219**
Maia-2 Descent	0.1512	± 0.0358
Maia-2 N. + Descent	0.1512	± 0.0358
Maia-2 FT + N. + Descent	0.1454	± 0.0388
Maia-2 MoE-LoRA N. + Descent	0.0438	± 0.0238
Maia-2 MCTS	0.1500	± 0.0387
Maia-2 N. + MCTS	0.1498	± 0.0385
Maia-2 FT + N. + MCTS	0.1399	± 0.0397
Maia-2 MoE-LoRA N. + MCTS	0.0434	± 0.0262

Table A.6: Overall Spatial Manifold Divergence (JSD_{spatial}) on full game records across model variants.

Model Variant	**Mean JSD$_{\text{spatial}}$**	**Std Dev**
Maia-2 Baseline	0.0265	± 0.0040
Maia-2 FT	0.0255	± 0.0044
Maia-2 Nucleus	0.0265	± 0.0040
Maia-2 MoE-LoRA	0.0069	± 0.0017
Maia-2 Descent	0.0265	± 0.0040
Maia-2 N. + Descent	0.0265	± 0.0040
Maia-2 FT + N. + Descent	0.0232	± 0.0033
Maia-2 MoE-LoRA N. + Descent	**0.0069**	± **0.0016**
Maia-2 MCTS	0.0257	± 0.0039
Maia-2 N. + MCTS	0.0257	± 0.0039
Maia-2 FT + N. + MCTS	0.0220	± 0.0036
Maia-2 MoE-LoRA N. + MCTS	0.0069	± 0.0017

A.2 Move Prediction Accuracy Breakdowns

Table A.7: Move Accuracy for Direct Policy Variants.

Player	**Maia-2 Baseline**	**Maia-2 FT**	**Maia-2 Nucleus**	**Maia-2 MoE-LoRA**
A. Alekhine	45.68%	45.46%	45.68%	**45.89%**
A. Beliavsky	43.94%	43.51%	43.94%	**46.21%**
A. Karpov	43.66%	43.22%	43.66%	**46.80%**
B. Larsen	**43.82%**	43.63%	**43.82%**	43.68%
G. Kasparov	45.32%	45.05%	45.32%	**46.07%**
J. Timman	44.39%	44.24%	44.39%	**45.88%**
L. Portisch	43.60%	43.19%	43.60%	**46.01%**
M. Tal	45.93%	45.80%	45.93%	**47.39%**
N. Short	43.99%	43.80%	43.99%	**45.55%**
T. Petrosian	43.02%	42.61%	43.02%	**43.39%**
U. Andersson	44.72%	44.27%	44.72%	**46.58%**
V. Ivanchuk	44.20%	44.06%	44.20%	**45.35%**
V. Korchnoi	43.79%	43.41%	43.79%	**45.33%**
V. Anand	45.16%	45.12%	45.16%	**47.98%**
Average	44.37%	44.10%	44.37%	**45.87%**

Table A.8: Move Accuracy for Descent Search Variants.

Player	**Maia-2 Descent**	**Maia-2 N. + Descent**	**Maia-2 FT + N. + Descent**	**Maia-2 MoE-LoRA N. + Descent**
A. Alekhine	45.68%	45.68%	44.73%	**46.19%**
A. Beliavsky	43.94%	43.94%	43.52%	**46.33%**
A. Karpov	43.67%	43.67%	42.65%	**46.61%**
B. Larsen	**43.81%**	**43.81%**	42.98%	43.41%
G. Kasparov	45.32%	45.32%	44.30%	**46.41%**
J. Timman	44.39%	44.39%	43.47%	**45.97%**
L. Portisch	43.61%	43.61%	43.17%	**46.33%**
M. Tal	45.93%	45.93%	44.85%	**47.48%**
N. Short	43.97%	43.98%	42.47%	**45.52%**
T. Petrosian	43.02%	43.02%	41.87%	**43.20%**
U. Andersson	44.71%	44.71%	43.41%	**46.62%**
V. Ivanchuk	44.19%	44.19%	43.33%	**45.34%**
V. Korchnoi	43.79%	43.79%	42.81%	**45.45%**
V. Anand	45.15%	45.15%	44.22%	**47.95%**
Average	44.37%	44.37%	43.41%	**45.91%**

Table A.9: Move Accuracy for MCTS Search Variants.

Player	**Maia-2 MCTS**	**Maia-2 N. + MCTS**	**Maia-2 FT + N. + MCTS**	**Maia-2 MoE-LoRA N. + MCTS**
A. Alekhine	49.02%	49.00%	48.34%	**49.32%**
A. Beliavsky	47.50%	47.43%	47.12%	**49.30%**
A. Karpov	46.86%	46.81%	46.04%	**49.81%**
B. Larsen	**47.26%**	47.23%	46.32%	46.82%
G. Kasparov	48.91%	48.86%	47.91%	**49.80%**
J. Timman	48.12%	48.06%	47.03%	**49.49%**
L. Portisch	46.88%	46.82%	46.45%	**49.47%**
M. Tal	49.63%	49.58%	48.00%	**50.81%**
N. Short	47.58%	47.52%	45.88%	**48.74%**
T. Petrosian	**46.01%**	45.95%	44.98%	45.83%
U. Andersson	47.50%	47.46%	46.01%	**49.25%**
V. Ivanchuk	47.65%	47.61%	46.79%	**48.72%**
V. Korchnoi	47.02%	46.96%	46.02%	**48.85%**
V. Anand	48.77%	48.70%	47.87%	**51.10%**
Average	47.76%	47.71%	46.77%	**49.09%**

A.3 Centipawn Error Breakdowns

Table A.10: Centipawn Error for Direct Policy Variants.

Player	**Maia-2 Baseline**	**Maia-2 FT**	**Maia-2 Nucleus**	**Maia-2 MoE-LoRA**
A. Alekhine	66.85	**62.16**	66.85	71.72
A. Beliavsky	58.25	**54.70**	58.25	61.43
A. Karpov	49.83	**47.53**	49.83	50.71
B. Larsen	62.85	**58.64**	62.85	66.11
G. Kasparov	54.42	**51.02**	54.42	60.49
J. Timman	57.89	**54.49**	57.89	59.98
L. Portisch	54.53	**50.92**	54.53	57.54
M. Tal	65.97	**61.40**	65.97	69.74
N. Short	62.27	**58.61**	62.27	65.90
T. Petrosian	41.22	**39.17**	41.22	47.19
U. Andersson	40.95	**38.33**	40.95	45.03
V. Ivanchuk	59.44	**55.57**	59.44	62.74
V. Korchnoi	54.73	**52.33**	54.73	55.27
V. Anand	54.98	**52.49**	54.98	59.03
Average	56.01	**52.67**	56.01	59.49

Table A.11: Centipawn Error for Descent Search Variants.

Player	**Maia-2 Descent**	**Maia-2 N. + Descent**	**Maia-2 FT + N. + Descent**	**Maia-2 MoE-LoRA N. + Descent**
A. Alekhine	66.74	66.17	**65.90**	72.78
A. Beliavsky	58.43	58.42	**56.64**	61.09
A. Karpov	49.58	49.58	**47.53**	50.74
B. Larsen	63.08	63.01	**62.17**	68.44
G. Kasparov	54.41	54.31	**54.25**	59.34
J. Timman	58.15	58.15	**56.28**	60.33
L. Portisch	54.68	54.68	**53.08**	57.36
M. Tal	66.14	66.05	**65.06**	71.29
N. Short	62.21	62.27	**61.68**	66.65
T. Petrosian	**41.21**	**41.21**	41.29	45.89
U. Andersson	41.12	41.04	**40.78**	44.32
V. Ivanchuk	59.59	59.58	**58.06**	59.64
V. Korchnoi	55.03	55.08	**54.54**	55.04
V. Anand	55.25	**55.12**	56.33	58.32
Average	56.12	56.05	**55.26**	59.38

Table A.12: Centipawn Error for MCTS Search Variants.

Player	**Maia-2 MCTS**	**Maia-2 N. + MCTS**	**Maia-2 FT + N. + MCTS**	**Maia-2 MoE-LoRA N. + MCTS**
A. Alekhine	27.87	25.92	**23.60**	33.29
A. Beliavsky	27.04	25.77	**24.85**	29.89
A. Karpov	22.72	22.18	**20.92**	24.18
B. Larsen	26.41	25.00	**22.21**	31.00
G. Kasparov	26.49	25.54	**23.47**	31.40
J. Timman	24.85	24.10	**21.97**	26.38
L. Portisch	25.76	24.09	**23.17**	27.23
M. Tal	27.75	27.08	**25.44**	33.50
N. Short	28.20	26.59	**26.07**	32.08
T. Petrosian	19.61	**18.17**	18.98	25.21
U. Andersson	21.05	19.87	**19.01**	22.88
V. Ivanchuk	27.27	25.45	**23.72**	27.04
V. Korchnoi	25.15	23.94	**22.63**	25.18
V. Anand	25.27	23.88	**23.69**	25.84
Average	25.39	24.11	**22.84**	28.22

A.4 Jensen-Shannon Divergence Breakdowns

Table A.13: Action-Level Jensen-Shannon Divergence (JSD_{action}) on common FENs for Direct Policy Variants.

Player	Maia-2 Baseline	Maia-2 FT	Maia-2 Nucleus	Maia-2 MoE-LoRA
A. Alekhine	0.4195	0.3981	0.4113	**0.2447**
A. Beliavsky	0.4277	0.4032	0.4205	**0.1782**
A. Karpov	0.4132	0.3913	0.4050	**0.1318**
B. Larsen	0.3945	0.3733	0.3875	**0.2223**
G. Kasparov	0.4323	0.4068	0.4244	**0.2228**
J. Timman	0.3967	0.3741	0.3899	**0.1638**
L. Portisch	0.4183	0.3867	0.4103	**0.2100**
M. Tal	0.3817	0.3493	0.3736	**0.1902**
N. Short	0.4367	0.4178	0.4299	**0.2430**
T. Petrosian	0.3896	0.3557	0.3824	**0.2064**
U. Andersson	0.4227	0.3946	0.4148	**0.1878**
V. Ivanchuk	0.3791	0.3558	0.3722	**0.1977**
V. Korchnoi	0.3769	0.3506	0.3709	**0.1905**
V. Anand	0.3983	0.3723	0.3905	**0.2050**
Average	0.4062	0.3807	0.3988	**0.1996**

Table A.14: Action-Level Jensen-Shannon Divergence (JSD_{action}) on common FENs for Descent Search Variants.

Player	Maia-2 Descent	Maia-2 N. + Descent	Maia-2 FT + N. + Descent	Maia-2 MoE-LoRA N. + Descent
A. Alekhine	0.7838	0.5094	0.5480	**0.3932**
A. Beliavsky	0.7934	0.5332	0.5632	**0.3375**
A. Karpov	0.7856	0.5185	0.5573	**0.2647**
B. Larsen	0.7582	0.4801	0.5107	**0.3685**
G. Kasparov	0.7855	0.5167	0.5558	**0.3889**
J. Timman	0.7697	0.4985	0.5242	**0.3382**
L. Portisch	0.7817	0.5129	0.5541	**0.3473**
M. Tal	0.7671	0.4865	0.5304	**0.3401**
N. Short	0.7800	0.5188	0.5525	**0.4108**
T. Petrosian	0.7656	0.4863	0.5422	**0.3607**
U. Andersson	0.7835	0.5133	0.5518	**0.3390**
V. Ivanchuk	0.7525	0.4700	0.5043	**0.3346**
V. Korchnoi	0.7533	0.4756	0.5071	**0.3264**
V. Anand	0.7782	0.5065	0.5436	**0.3665**
Average	0.7741	0.5019	0.5389	**0.3512**

Table A.15: Action-Level Jensen-Shannon Divergence (JSD_{action}) on common FENs for MCTS Search Variants.

Player	Maia-2 MCTS	Maia-2 N. + MCTS	Maia-2 FT + N. + MCTS	Maia-2 MoE-LoRA N. + MCTS
A. Alekhine	0.3988	0.3990	0.3689	**0.2182**
A. Beliavsky	0.4029	0.4036	0.3785	**0.1728**
A. Karpov	0.3885	0.3887	0.3656	**0.1573**
B. Larsen	0.3750	0.3755	0.3385	**0.2076**
G. Kasparov	0.4098	0.4104	0.3800	**0.1974**
J. Timman	0.3703	0.3710	0.3474	**0.1744**
L. Portisch	0.3927	0.3932	0.3618	**0.1791**
M. Tal	0.3570	0.3577	0.3186	**0.1529**
N. Short	0.4169	0.4173	0.3981	**0.2349**
T. Petrosian	0.3680	0.3684	0.3543	**0.1954**
U. Andersson	0.3979	0.3983	0.3710	**0.1775**
V. Ivanchuk	0.3559	0.3565	0.3104	**0.1813**
V. Korchnoi	0.3550	0.3556	0.3286	**0.1623**
V. Anand	0.3722	0.3726	0.3408	**0.1945**
Average	0.3829	0.3834	0.3544	**0.1861**

Table A.16: Action-Level Jensen-Shannon Divergence (JSD_{action}) on full game records for Direct Policy Variants.

Player	Maia-2 Baseline	Maia-2 FT	Maia-2 Nucleus	Maia-2 MoE-LoRA
A. Alekhine	0.5014	0.5035	**0.4934**	0.5221
A. Beliavsky	0.5256	0.5279	**0.5182**	0.5438
A. Karpov	0.5269	0.5315	**0.5194**	0.5409
B. Larsen	0.5220	0.5243	**0.5145**	0.5360
G. Kasparov	0.5153	0.5193	**0.5073**	0.5381
J. Timman	0.5173	0.5201	**0.5098**	0.5331
L. Portisch	0.5218	0.5255	**0.5143**	0.5390
M. Tal	0.5089	0.5117	**0.5013**	0.5271
N. Short	0.5252	0.5285	**0.5176**	0.5417
T. Petrosian	0.5316	0.5378	**0.5242**	0.5519
U. Andersson	0.5211	0.5251	**0.5135**	0.5374
V. Ivanchuk	0.5208	0.5247	**0.5135**	0.5339
V. Korchnoi	0.5241	0.5287	**0.5168**	0.5367
V. Anand	0.5154	0.5194	**0.5076**	0.5287
Average	0.5198	0.5234	**0.5122**	0.5365

Table A.17: Action-Level Jensen-Shannon Divergence (JSD_{action}) on full game records for Descent Search Variants.

Player	**Maia-2 Descent**	**Maia-2 N. + Descent**	**Maia-2 FT + N. + Descent**	**Maia-2 MoE-LoRA N. + Descent**
A. Alekhine	0.8412	**0.5800**	0.6904	0.5877
A. Beliavsky	0.8417	**0.5977**	0.7029	0.6077
A. Karpov	0.8421	**0.5998**	0.7064	0.6085
B. Larsen	0.8391	**0.5948**	0.7014	0.6054
G. Kasparov	0.8472	**0.5929**	0.7076	0.6021
J. Timman	0.8392	**0.5892**	0.6976	0.6002
L. Portisch	0.8441	**0.5948**	0.7067	0.6045
M. Tal	0.8426	**0.5833**	0.6957	0.5927
N. Short	0.8407	**0.6001**	0.7054	0.6102
T. Petrosian	0.8525	**0.6075**	0.7195	0.6151
U. Andersson	0.8421	**0.5950**	0.7018	0.6002
V. Ivanchuk	0.8386	**0.5924**	0.6989	0.6033
V. Korchnoi	0.8421	**0.5970**	0.7053	0.6092
V. Anand	0.8404	**0.5916**	0.6988	0.5999
Average	0.8424	**0.5940**	0.7027	0.6033

Table A.18: Action-Level Jensen-Shannon Divergence (JSD_{action}) on full game records for MCTS Search Variants.

Player	**Maia-2 MCTS**	**Maia-2 N. + MCTS**	**Maia-2 FT + N. + MCTS**	**Maia-2 MoE-LoRA N. + MCTS**
A. Alekhine	0.4594	**0.4587**	0.4956	0.4837
A. Beliavsky	0.4838	**0.4834**	0.5200	0.5048
A. Karpov	0.4871	**0.4869**	0.5245	0.5005
B. Larsen	0.4818	**0.4814**	0.5185	0.4985
G. Kasparov	0.4719	**0.4715**	0.5120	0.4932
J. Timman	0.4757	**0.4752**	0.5113	0.4905
L. Portisch	0.4816	**0.4812**	0.5198	0.4976
M. Tal	0.4662	**0.4657**	0.5028	0.4847
N. Short	0.4841	**0.4836**	0.5215	0.5018
T. Petrosian	0.4944	**0.4943**	0.5328	0.5203
U. Andersson	0.4842	**0.4838**	0.5232	0.4975
V. Ivanchuk	0.4803	**0.4798**	0.5159	0.4929
V. Korchnoi	0.4840	**0.4837**	0.5218	0.4944
V. Anand	0.4736	**0.4731**	0.5130	0.4872
Average	0.4792	**0.4787**	0.5166	0.4963

A.5 Stylistic Manifold Alignment Breakdowns

Table A.19: Spatial Manifold Divergence (JSD_{spatial}) on common FENs for Direct Policy Variants.

Player	Maia-2 Baseline	Maia-2 FT	Maia-2 Nucleus	Maia-2 MoE-LoRA
A. Alekhine	0.2038	0.2277	0.2038	**0.0904**
A. Beliavsky	0.1119	0.1594	0.1119	**0.0294**
A. Karpov	0.1432	0.1738	0.1432	**0.0397**
B. Larsen	0.1643	0.1159	0.1643	**0.0750**
G. Kasparov	0.1821	0.1511	0.1821	**0.0675**
J. Timman	0.1297	0.1415	0.1297	**0.0514**
L. Portisch	0.1450	0.1571	0.1450	**0.0335**
M. Tal	0.1505	0.1567	0.1505	**0.0242**
N. Short	0.2032	0.2211	0.2032	**0.0116**
T. Petrosian	0.1056	0.1207	0.1056	**0.0268**
U. Andersson	0.0788	0.0908	0.0788	**0.0291**
V. Ivanchuk	0.1454	0.1314	0.1454	**0.0466**
V. Korchnoi	0.1604	0.1114	0.1604	**0.0279**
V. Anand	0.1935	0.2014	0.1935	**0.0203**
Average	0.1512	0.1543	0.1512	**0.0409**

Table A.20: Spatial Manifold Divergence (JSD_{spatial}) on common FENs for Descent Search Variants.

Player	Maia-2 Descent	Maia-2 N. + Descent	Maia-2 FT + N. + Descent	Maia-2 MoE-LoRA N. + Descent
A. Alekhine	0.2038	0.2038	0.2191	**0.0782**
A. Beliavsky	0.1119	0.1119	0.1125	**0.0242**
A. Karpov	0.1432	0.1432	0.1500	**0.0396**
B. Larsen	0.1643	0.1643	0.1547	**0.0888**
G. Kasparov	0.1821	0.1821	0.1654	**0.0684**
J. Timman	0.1297	0.1297	0.1465	**0.0526**
L. Portisch	0.1450	0.1450	0.1110	**0.0141**
M. Tal	0.1505	0.1505	0.1354	**0.0234**
N. Short	0.2032	0.2032	0.2076	**0.0122**
T. Petrosian	0.1056	0.1056	0.1602	**0.0651**
U. Andersson	0.0788	0.0788	0.0792	**0.0281**
V. Ivanchuk	0.1454	0.1454	0.1232	**0.0474**
V. Korchnoi	0.1604	0.1604	0.0931	**0.0514**
V. Anand	0.1935	0.1935	0.1771	**0.0204**
Average	0.1512	0.1512	0.1454	**0.0438**

Table A.21: Spatial Manifold Divergence (JSD_{spatial}) on common FENs for MCTS Search Variants.

Player	Maia-2 MCTS	Maia-2 N. + MCTS	Maia-2 FT + N. + MCTS	Maia-2 MoE-LoRA N. + MCTS
A. Alekhine	0.2066	0.2064	0.2128	**0.0833**
A. Beliavsky	0.1085	0.1082	0.1195	**0.0311**
A. Karpov	0.1420	0.1420	0.1207	**0.0405**
B. Larsen	0.1749	0.1735	0.1471	**0.1068**
G. Kasparov	0.1755	0.1732	0.1700	**0.0737**
J. Timman	0.1118	0.1119	0.1603	**0.0533**
L. Portisch	0.1441	0.1455	0.1200	**0.0330**
M. Tal	0.1551	0.1546	0.1460	**0.0238**
N. Short	0.2092	0.2092	0.2083	**0.0116**
T. Petrosian	0.1041	0.1041	0.1371	**0.0311**
U. Andersson	0.0740	0.0739	0.0714	**0.0290**
V. Ivanchuk	0.1436	0.1434	0.1550	**0.0462**
V. Korchnoi	0.1609	0.1611	0.0785	**0.0274**
V. Anand	0.1902	0.1900	0.1121	**0.0165**
Average	0.1500	0.1498	0.1399	**0.0434**

Table A.22: Spatial Manifold Divergence (JSD_{spatial}) on full game records for Direct Policy Variants.

Player	Maia-2 Baseline	Maia-2 FT	Maia-2 Nucleus	Maia-2 MoE-LoRA
A. Alekhine	0.0298	0.0292	0.0298	**0.0095**
A. Beliavsky	0.0310	0.0331	0.0310	**0.0046**
A. Karpov	0.0296	0.0277	0.0296	**0.0066**
B. Larsen	0.0163	0.0157	0.0163	**0.0094**
G. Kasparov	0.0296	0.0281	0.0296	**0.0075**
J. Timman	0.0251	0.0235	0.0251	**0.0077**
L. Portisch	0.0298	0.0298	0.0298	**0.0067**
M. Tal	0.0273	0.0255	0.0273	**0.0063**
N. Short	0.0279	0.0270	0.0279	**0.0055**
T. Petrosian	0.0261	0.0254	0.0261	**0.0094**
U. Andersson	0.0204	0.0188	0.0204	**0.0051**
V. Ivanchuk	0.0254	0.0236	0.0254	**0.0068**
V. Korchnoi	0.0237	0.0224	0.0237	**0.0079**
V. Anand	0.0293	0.0275	0.0293	**0.0041**
Average	0.0265	0.0255	0.0265	**0.0069**

Table A.23: Spatial Manifold Divergence (JSD_{spatial}) on full game records for Descent Search Variants.

Player	Maia-2 Descent	Maia-2 N. + Descent	Maia-2 FT + N. + Descent	Maia-2 MoE-LoRA N. + Descent
A. Alekhine	0.0298	0.0298	0.0267	**0.0093**
A. Beliavsky	0.0310	0.0310	0.0255	**0.0049**
A. Karpov	0.0296	0.0296	0.0251	**0.0066**
B. Larsen	0.0163	0.0163	0.0153	**0.0090**
G. Kasparov	0.0296	0.0296	0.0242	**0.0074**
J. Timman	0.0250	0.0250	0.0212	**0.0076**
L. Portisch	0.0298	0.0298	0.0250	**0.0066**
M. Tal	0.0273	0.0273	0.0264	**0.0059**
N. Short	0.0280	0.0280	0.0265	**0.0051**
T. Petrosian	0.0261	0.0261	0.0244	**0.0089**
U. Andersson	0.0204	0.0204	0.0189	**0.0052**
V. Ivanchuk	0.0253	0.0253	0.0195	**0.0070**
V. Korchnoi	0.0236	0.0236	0.0221	**0.0082**
V. Anand	0.0294	0.0294	0.0243	**0.0043**
Average	0.0265	0.0265	0.0232	**0.0069**

Table A.24: Spatial Manifold Divergence (JSD_{spatial}) on full game records for MCTS Search Variants.

Player	Maia-2 MCTS	Maia-2 N. + MCTS	Maia-2 FT + N. + MCTS	Maia-2 MoE-LoRA N. + MCTS
A. Alekhine	0.0301	0.0301	0.0279	**0.0093**
A. Beliavsky	0.0304	0.0304	0.0215	**0.0047**
A. Karpov	0.0286	0.0285	0.0222	**0.0065**
B. Larsen	0.0160	0.0160	0.0171	**0.0089**
G. Kasparov	0.0288	0.0287	0.0258	**0.0079**
J. Timman	0.0239	0.0240	0.0215	**0.0068**
L. Portisch	0.0296	0.0297	0.0215	**0.0068**
M. Tal	0.0257	0.0257	0.0275	**0.0056**
N. Short	0.0254	0.0255	0.0270	**0.0048**
T. Petrosian	0.0261	0.0261	0.0207	**0.0091**
U. Andersson	0.0199	0.0199	0.0190	**0.0057**
V. Ivanchuk	0.0244	0.0244	0.0202	**0.0081**
V. Korchnoi	0.0235	0.0235	0.0197	**0.0083**
V. Anand	0.0272	0.0272	0.0167	**0.0041**
Average	0.0257	0.0257	0.0220	**0.0069**