Novelty Assessment Report

Paper: Quasi-Equivariant Metanetworks

PDF URL: https://openreview.net/pdf?id=XMiDpi2mWY

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-01

Abstract

Metanetworks are neural architectures designed to operate directly on pretrained weights to perform downstream tasks. However, the parameter space serves only as a proxy for the underlying function class, and the parameter-function mapping is inherently non-injective: distinct parameter configurations may yield identical input-output behaviors. As a result, metanetworks that rely solely on raw parameters risk overlooking the intrinsic symmetries of the architecture. Reasoning about functional identity is therefore essential for effective metanetwork design, motivating the development of equivariant metanetworks, which incorporate equivariance principles to respect architectural symmetries. Existing approaches, however, typically enforce strict equivariance, which imposes rigid constraints and often leads to sparse and less expressive models. To address this limitation, we introduce the novel concept of quasi-equivariance, which allows metanetworks to move beyond the rigidity of strict equivariance while still preserving functional identity. We lay down a principled basis for this framework and demonstrate its broad applicability across diverse neural architectures, including feedforward, convolutional, and transformer networks. Through empirical evaluation, we show that quasi-equivariant metanetworks achieve good trade-offs between symmetry preservation and representational expressivity. These findings advance the theoretical understanding of weight-space learning and provide a principled foundation for the design of more expressive and functionally robust metanetworks.

Core Task Landscape

This paper addresses: **Designing Metanetworks that Operate on Neural Network Weights**

A total of **49 papers** were analyzed and organized into a taxonomy with **33 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Weight-Space Representation Learning and Transformation**
- **Weight Generation and Synthesis**
- **Equivariant and Permutation-Aware Metanetworks**
- **Meta-Learning for Weight Initialization and Architecture**
- **Metanetworks for Network Pruning and Compression**
- **Metanetworks for Model Editing and Adaptation**
- **Theoretical Foundations of Weight-Space Symmetries**
- **Specialized and Domain-Specific Applications**
- **Peripheral and Loosely Related Work**

Complete Taxonomy Tree

- Designing Metanetworks that Operate on Neural Network Weights Survey Taxonomy
- Weight-Space Representation Learning and Transformation
  - Weight-Space Autoencoders and Embeddings (2 papers)
  - [4] Hyper-Representations as Generative Models: Sampling Unseen Neural Network Weights (SchÃ¼rholt, 2022) View paper
  - [48] Metanetwork: A novel approach to interpreting ANNs (R Takatsuki, n.d.) View paper
  - Weight Alignment and Symmetry Handling (2 papers)
  - [2] Equivariant deep weight space alignment (Navon, 2023) View paper
  - [28] Symmetry-Aware Graph Metanetwork Autoencoders: Model Merging through Parameter Canonicalization (Odysseas Boufalis, 2025) View paper
  - Compressed Weight-Space Operations (1 papers)
  - [34] Training and Generating Neural Networks in Compressed Weight Space (Irie, 2021) View paper
- Weight Generation and Synthesis
  - Hypernetworks and Direct Weight Generation (3 papers)
  - [12] Weightnet: Revisiting the design space of weight networks (Ningning Ma, 2020) View paper
  - [22] Generating Synaptic Weights from Task Specifications: A Computational Paradigm for Neural Network Initialization (Culaj, 2025) View paper
  - [43] Generating Neural Networks with Neural Networks (Deutsch, 2022) View paper
  - Text-Conditioned and Task-Conditioned Weight Generation (1 papers)
  - [17] Text2Weight: Bridging Natural Language and Neural Network Weight Spaces (Bowen Tian, 2025) View paper
- Equivariant and Permutation-Aware Metanetworks
  - Strict Equivariant Metanetwork Architectures (4 papers)
  - [7] Scale equivariant graph metanetworks (Giorgos Bouritsas, 2024) View paper
  - [16] Universal Neural Functionals (Chelsea Finn, 2024) View paper
  - [18] GL Equivariant Metanetworks for Learning on Low Rank Weight Spaces (T Putterman, 2025) View paper

- [35] Symmetry-Aware Fully-Amortized Optimization with Scale Equivariant Graph Metanetworks (Bart Kuipers, 2025) View paper
  - Quasi-Equivariant and Relaxed Symmetry Metanetworks ★ (1 papers)
  - [0] Quasi-Equivariant Metanetworks (Anon et al., 2026) View paper
  - Graph-Based Metanetworks for Diverse Architectures (1 papers)
  - [36] Graph Metanetworks for Processing Diverse Neural Architectures (Lim, 2023) View paper
  - Neural Functional Transformers (1 papers)
  - [41] Neural Functional Transformers (Zhou, 2023) View paper
- Meta-Learning for Weight Initialization and Architecture
  - Meta-Learning for Model Initialization (1 papers)
  - [46] General-Purpose In-Context Learning by Meta-Learning Transformers (Kirsch, 2022) View paper
  - Joint Architecture and Meta-Weight Optimization (1 papers)
  - [29] Learning to learn by jointly optimizing neural architecture and weights (Yadong Ding, 2022) View paper
  - Meta-Learning for Architecture Search (4 papers)
  - [9] Rapid model architecture adaption for meta-learning (Zhao, 2022) View paper
  - [25] EMAS: Efficient Meta Architecture Search for Few-Shot Learning (Dongkai Liu, 2022) View paper
  - [26] Hardware-adaptive efficient latency prediction for nas via meta-learning (Ha-Yeon Lee, 2021) View paper
  - [27] Help: Hardware-adaptive efficient latency prediction for nas via meta-learning (Lee HaYeon, 2021) View paper
- Metanetworks for Network Pruning and Compression
  - Meta-Learning for Automatic Channel Pruning (1 papers)
  - [1] Metapruning: Meta learning for automatic neural network channel pruning (Zechun Liu, 2019) View paper
  - Graph Metanetworks for Universal Pruning (2 papers)
  - [14] Meta Pruning via Graph Metanetworks : A Universal Meta Learning Framework for Network Pruning (Liu Ye-wei, 2025) View paper
  - Meta-Learning for Sparse Structure Optimization (1 papers)
  - [6] Meta-Sparsity: Learning Optimal Sparse Structures in Multi-task Networks through Meta-learning (Upadhyay, 2025) View paper
- Metanetworks for Model Editing and Adaptation
  - Regulatory Metanetworks for Requirement Compliance (1 papers)
  - [33] Metanetworks as Regulatory Operators: Learning to Edit for Requirement Compliance (Ioannis Kalogeropoulos, 2025) View paper
  - Checkpoint-Based Hyperparameter Optimization (1 papers)
  - [8] Improving Hyperparameter Optimization with Checkpointed Model Weights (Nikhil Mehta, 2024) View paper
  - Recursive Transformers with Conditional Modulation (1 papers)
  - [37] Improving Recursive Transformers with Mixture of LoRAs (Mohammadmahdi Nouriborji, 2025) View paper
- Theoretical Foundations of Weight-Space Symmetries
  - Algebraic and Geometric Symmetry Analysis (1 papers)
  - [40] On the geometry of feedforward neural network error surfaces (An Mei Chen, 1993) View paper
  - Genetic Encoding and Symmetry Elimination (1 papers)
  - [38] Non-redundant genetic coding of neural networks (D. Thierens, 1996) View paper
- Specialized and Domain-Specific Applications
  - Physics-Informed Neural Networks with Meta-Learning (1 papers)
  - [11] Accelerating physics-informed neural network based 1D arc simulation by meta learning (L. Zhong, 2023) View paper
  - Sample Weighting via Meta-Networks (1 papers)
  - [31] FMW-Net: a first-order meta-weight-net approach for sample weighting (Yubo Zhou, 2025) View paper
  - On-Device Neural Architecture Search (1 papers)
  - [13] On-NAS: On-device neural architecture search on memory-constrained intelligent embedded systems (Bosung Kim, 2023) View paper
  - Network Transformation for Efficient Architecture Search (1 papers)
  - [49] Efficient Architecture Search by Network Transformation (Han Cai, 2017) View paper
  - Hypergraph Neural Networks with Meta-Learning Attention (1 papers)
  - [20] Overlap-aware meta-learning attention to enhance hypergraph neural networks for node classification (Yang Mu-rong, 2025) View paper
  - Recursive Meta-Networks for Self-Modification (1 papers)
  - [30] A neural network that embeds its own meta-levels (J Schmidhuber, 1993) View paper
  - Meta-Learning for Neural Network Optimization (1 papers)
  - [32] Meta-learning approach to neural network optimization (Pavel KordÃk, 2010) View paper
- Peripheral and Loosely Related Work
  - Network Dynamics and Learning Theory (1 papers)
  - [19] Network Dynamics-Based Framework for Understanding Deep Neural Networks (Lin Yu-Chen, 2025) View paper
  - Weight-Space Transformations in Classical Networks (3 papers)
  - [21] Exploring Local Transformation Shared Weights in Convolutional Neural Networks (Rohan Ghosh, 2019) View paper
  - [44] A connectionist approach to speech recognition (Yoshua Bengio, 1993) View paper
  - [47] Transformations of sigmaâpi nets: obtaining reflected functions by reflecting weight matrices (R. Neville, 2002) View paper
  - Element-Wise Multiplicative Operators (1 papers)
  - [39] Element-Wise Multiplicative Operators in Vision, Language, and Multimodal Learning (Gurpreet Inayatullah, 2025) View paper
  - Modularity and Meta-Architectures in Deep Learning (1 papers)
  - [5] Modularity in deep learning (Sun, 2023) View paper
  - Domain-Specific Applications Unrelated to Metanetworks (5 papers)
  - [3] Image-based plant disease identification by deep learning meta-architectures (Muhammad Hammad Saleem, 2020) View paper
  - [10] Analysis of weighted networks (M.E.J. Newman, 2004) View paper
  - [23] A deep patch network with spatiotemporal meta-parameter learning for soft sensor modeling of industrial processes (Xudong Shi, 2025) View paper
  - [24] Detection of food allergy using deep learning (Abdul Majid Soomro, 2023) View paper

- [42] Deep E-Learning RecommendNet: An Acute E-Learning Recommendation System with Meta-Heuristic-Based Hybrid Deep Learning Architecture (Pradnya Vaibhav Kulkarni, 2022) View paper
- General Machine Learning and AI Foundations (1 papers)
- [45] Engineering artificial systems with natural intelligence (Raghavan, 2023) View paper

Narrative

Core task: Designing metanetworks that operate on neural network weights. This emerging field explores how one neural network can process, transform, or generate the weights of another network, treating weight tensors as structured data rather than opaque parameters. The taxonomy reveals several major branches: Weight-Space Representation Learning focuses on encoding and transforming existing weights (e.g., Deep Weight Alignment[2], Checkpointed Model Weights[8]); Weight Generation and Synthesis addresses creating new network parameters from scratch or from high-level specifications (e.g., Text2Weight[17], Generating Synaptic Weights[22]); Equivariant and Permutation-Aware Metanetworks emphasizes respecting the inherent symmetries of weight spaces, including strict equivariance (Universal Neural Functionals[16], GL Equivariant Metanetworks[18]) and relaxed approaches; Meta-Learning branches explore using metanetworks for initialization and architecture search (Rapid Architecture Adaption[9]); and Pruning and Compression apply metanetworks to model efficiency (Metapruning[1], Graph Metanetworks Pruning[14]). Additional branches cover model editing, theoretical symmetry foundations, and domain-specific applications.

A central tension runs through the field between strict equivariance—which guarantees that metanetwork outputs respect weight-space permutation symmetries—and more flexible, quasi-equivariant designs that trade theoretical guarantees for practical expressiveness or computational efficiency. Works like Scale Equivariant Metanetworks[7] and Symmetry-Aware Autoencoders[28] pursue rigorous symmetry preservation, while Quasi-Equivariant Metanetworks[0] explores relaxed symmetry constraints that maintain useful inductive biases without full equivariance. This positioning reflects a broader question: when operating on weights, how much structure should be baked into the metanetwork architecture versus learned from data? The original paper sits within the quasi-equivariant cluster, contrasting with fully equivariant approaches by allowing controlled symmetry violations that may enhance flexibility for tasks where approximate invariance suffices, such as weight editing or cross-architecture transfer where strict permutation equivariance may be overly restrictive.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

The original leaf focuses on metanetworks that intentionally relax strict equivariance constraints to gain expressivity while maintaining some symmetry properties. Its siblings represent alternative architectural approaches: graph-based methods that encode weights as graphs, transformer-based attention mechanisms with permutation equivariance, and strictly equivariant architectures that enforce exact symmetry preservation. The key distinction is the trade-off between strict mathematical guarantees (strict equivariance) versus practical flexibility (quasi-equivariance).

Similarities: - All subtopics address the fundamental challenge of processing neural network weights while respecting their inherent symmetries (permutation, scaling) - Each approach aims to create metanetworks that can generalize across different weight configurations - All methods must handle the unique structure of weight spaces compared to standard data domains

Differences: - Quasi-Equivariant approaches explicitly trade strict symmetry for expressivity, while Strict Equivariant methods maintain exact mathematical guarantees - Graph-Based methods use graph neural networks as the primary architectural choice, encoding weights as graph structures - Neural Functional Transformers leverage attention mechanisms and transformer architectures, while the original leaf is architecture-agnostic regarding the relaxation strategy - Strict Equivariant architectures prioritize theoretical correctness, whereas Quasi-Equivariant methods prioritize practical performance and representational capacity

Suggested Search Directions: - Empirical comparisons between quasi-equivariant and strictly equivariant methods on benchmark tasks - Theoretical analysis of what expressivity is gained by relaxing equivariance constraints - Hybrid approaches that combine graph-based or transformer architectures with relaxed symmetry constraints

Sibling Subtopics

- **Graph-Based Metanetworks for Diverse Architectures** (leaves: 1, papers: 1)
- Scope: Graph neural network approaches that process weights of diverse neural architectures by encoding them as graphs.
- Exclude: Excludes non-graph-based equivariant methods or fixed-architecture metanetworks; see Strict Equivariant Metanetwork Architectures.
- **Neural Functional Transformers** (leaves: 1, papers: 1)
- Scope: Transformer-based architectures that operate on neural network weights using attention mechanisms with permutation equivariance.
- Exclude: Excludes non-transformer equivariant architectures; see Strict Equivariant Metanetwork Architectures or Graph-Based Metanetworks.
- **Strict Equivariant Metanetwork Architectures** (leaves: 1, papers: 4)
- Scope: Metanetworks enforcing strict equivariance to permutation and scaling transformations in weight space.
- Exclude: Excludes quasi-equivariant or relaxed symmetry approaches; see Quasi-Equivariant and Relaxed Symmetry Metanetworks.

Contributions Analysis

Overall novelty summary. The paper introduces a quasi-equivariance framework for metanetworks that operate on neural network weights, relaxing strict equivariance constraints to balance symmetry preservation with representational expressivity. According to the taxonomy, this work occupies the 'Quasi-Equivariant and Relaxed Symmetry Metanetworks' leaf under the broader 'Equivariant and Permutation-Aware Metanetworks' branch. Notably, this leaf contains only the original paper itself—no sibling papers are present. This positioning suggests the paper addresses a relatively sparse research direction within the metanetwork landscape, where most prior work has focused on either strict equivariance or non-equivariant approaches.

The taxonomy reveals that the paper's immediate neighbors are in the 'Strict Equivariant Metanetwork Architectures' leaf, which contains four papers enforcing rigorous permutation and scaling symmetries. Adjacent leaves include 'Graph-Based Metanetworks for Diverse Architectures' and 'Neural Functional Transformers', both pursuing equivariance through different architectural paradigms. The scope note for the original paper's leaf explicitly excludes strictly equivariant architectures, positioning quasi-equivariance as a distinct middle ground between full symmetry enforcement and unconstrained weight-space operations. This structural context suggests the paper carves out conceptual space between established strict-equivariance methods and general weight-space learning approaches.

Among 29 candidates examined across three contributions, the theoretical foundation contribution shows the most substantial prior work overlap: 5 of 10 examined candidates appear refutable, indicating that connecting symmetry groups to functional equivalence has been explored in related contexts. In contrast, the quasi-equivariance framework itself and the general construction method show no clear refutations among their respective 10 and 9 examined candidates. This pattern suggests that while the underlying theoretical machinery may build on established symmetry analysis, the specific quasi-equivariant formulation and its practical instantiation represent less-explored territory within the limited search scope.

Based on the top-29 semantic matches examined, the work appears to occupy a genuinely sparse niche—being the sole occupant of its taxonomy leaf—though the theoretical underpinnings connect to a more developed literature on weight-space symmetries. The limited search scope means we cannot definitively assess novelty against the entire field, but the structural isolation within the taxonomy and the contribution-level statistics suggest the quasi-equivariance concept itself is relatively unexplored, even if it builds on established symmetry theory.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Quasi-equivariance framework for metanetworks

Description: The authors propose quasi-equivariance as a relaxation of strict equivariance that maintains functional identity while providing greater representational flexibility. This framework enables metanetworks to preserve functional equivalence classes without the rigid constraints imposed by strict equivariance.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Relaxing equivariance constraints with non-stationary continuous filters

URL: View paper

Brief Assessment

Non-Stationary Filters[50] focuses on relaxing equivariance constraints in standard neural network layers (CNNs, transformers) through non-stationary continuous filters, not on metanetworks that operate on pretrained weights. The original paper's quasi-equivariance framework specifically addresses functional equivalence classes in weight-space learning for metanetworks, which is a fundamentally different problem domain.

---

2. On the scale invariance in state of the art CNNs trained on ImageNet

URL: View paper

Brief Assessment

Scale Invariance CNNs[55] focuses on analyzing scale invariance in CNNs trained on ImageNet for computer vision tasks, not on relaxing equivariance constraints in metanetworks that operate on neural network weights.

---

3. Sharp minima can generalize for deep nets

URL: View paper

Brief Assessment

Sharp Minima Generalization[57] focuses on the geometry of loss surfaces and parameter space transformations in deep networks, not on equivariance frameworks for metanetworks. The papers address fundamentally different problems with no overlap in methodology or objectives.

---

4. Approximately equivariant graph networks

URL: View paper

Brief Assessment

Approximately Equivariant[53] focuses on relaxing equivariance for graph neural networks on fixed graph domains with active symmetries, not metanetworks operating on neural network weights. The candidate addresses approximate symmetries via graph coarsening for learning on graphs, while the original addresses functional equivalence classes in weight spaces of neural networks.

---

5. Weakly connected neural networks

URL: View paper

Brief Assessment

Weakly Connected Networks[58] focuses on topological equivalence and synaptic connection weakness in neural network dynamics, not on relaxing equivariance constraints in metanetworks while preserving functional identity. The candidate addresses fundamentally different aspects of neural networks.

---

6. Equivariance-aware architectural optimization of neural networks

URL: View paper

Brief Assessment

Equivariance-Aware Optimization[54] focuses on architectural optimization of neural networks through relaxation morphisms and mixed layers for tasks like image classification, not on metanetworks that operate on pretrained weights. The candidate addresses approximate equivariance in standard networks, while the original addresses functional equivalence in weight-space learning for metanetworks.

---

7. Self-supervised image denoising with downsampled invariance loss and conditional blind-spot network

URL: View paper

Brief Assessment

Downsampled Invariance Loss[56] focuses on self-supervised image denoising using conditional blind-spot networks and downsampled invariance loss for handling spatially correlated noise in images. This is fundamentally different from the original paper's quasi-equivariance framework for metanetworks, which addresses relaxing equivariance constraints in neural networks that operate on pretrained weights while preserving functional identity.

---

8. AV-NAS: Audio-Visual Multi-Level Semantic Neural Architecture Search for Video Hashing

URL: View paper

Brief Assessment

AV-NAS[51] focuses on neural architecture search for audio-visual video hashing, not on relaxing equivariance constraints in metanetworks or preserving functional identity in weight-space learning.

---

9. Relaxed Equivariant Graph Neural Networks

URL: View paper

Brief Assessment

Relaxed Equivariant[59] focuses on relaxing E(3) equivariance in graph neural networks for physical systems with symmetry breaking, not on metanetworks operating on neural network weights. The technical domains and applications are fundamentally different.

---

10. Symmetry breaking and equivariant neural networks

URL: View paper

Brief Assessment

Symmetry Breaking Equivariance[52] focuses on relaxed equivariance for general neural networks handling symmetry breaking in physics and graph tasks, not specifically for metanetworks operating on pretrained weights. The candidate addresses a different problem domain (symmetry breaking in data samples) rather than functional equivalence in weight spaces.

Contribution 2: Principled theoretical foundation connecting symmetry groups to functional equivalence

Description: The work establishes a formal theoretical foundation by analyzing parameter spaces, characterizing symmetry groups, and introducing the notion of maximal symmetry groups. This provides a principled connection between group-theoretic symmetries and functional equivalence in neural networks.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Universal approximations of invariant maps by neural networks

URL: View paper

Brief Assessment

Universal Invariant Approximations[69] focuses on universal approximation theorems for invariant/equivariant neural network maps with respect to group representations, not on characterizing functional equivalence classes in parameter spaces or establishing maximal symmetry groups as done in the original paper.

---

2. A symmetry-aware exploration of bayesian neural network posteriors

URL: View paper

Prior Art Analysis

Bayesian Posteriors Symmetry[76] demonstrates that prior work exists on characterizing symmetry groups and their connection to functional equivalence in neural networks. The candidate paper explicitly analyzes parameter spaces, characterizes symmetry groups (including permutation and scaling symmetries), and establishes formal connections between these symmetries and functional equivalence. The candidate provides mathematical definitions and propositions showing how symmetries create functionally equivalent parameter configurations, directly addressing the same theoretical foundation claimed as novel by the original paper.

Evidence

Evidence 1 - **Rationale**: Both papers address the same core concept: multiple parameter configurations yielding identical functions. The candidate provides formal mathematical definitions for this phenomenon. - **Original**: a major challenge in designing metanetworks lies on how to capture functional equivalence - the fact that multiple distinct parameter configurations can realize the same input-output function - **Candidate**: weight-space symmetries transform the parameters of the neural networks while keeping the networks functionally invariant; in other words, definition 3.1. let fω be a neural network of parameters ω taking n-dimensional vectors as inputs. we say that the transformation t modifying ω is a weight-space...

Evidence 2 - **Rationale**: The candidate paper characterizes symmetry groups (permutation and scaling) and establishes their connection to functional equivalence through formal propositions, demonstrating prior theoretical work on this foundation. - **Original**: we examine the parameter space of a parameterized function, characterize its associated symmetry group, and introduce the formal notion of maximality within symmetry groups, establishing a direct connection to functional equivalence. - **Candidate**: proposition 1. define fω a neural network and f˜ω its corresponding identifiable model - a network transformed for having sorted unit-normed neurons. let us also denote π and λ, the sets of permutation sets and scaling sets, respectively, and ˜ω the random variable of the sorted weights with unit no...

---

3. Probabilistic symmetries and invariant neural networks

URL: View paper

Prior Art Analysis

Probabilistic Symmetries[72] demonstrates that prior work established formal theoretical foundations connecting group-theoretic symmetries to functional equivalence in neural networks. The candidate paper presents a comprehensive framework analyzing parameter spaces, characterizing symmetry groups, and introducing maximal symmetry groups to establish connections between probabilistic and functional symmetry. This work predates the original paper and provides the theoretical machinery (maximal invariants, adequacy, sufficency) that the original paper claims as novel contributions.

Evidence

Evidence 1 - **Rationale**: The candidate establishes general theoretical results (Theorem 7) characterizing all invariant neural network architectures under compact group actions, providing the principled theoretical foundation that connects group symmetries to functional representations across diverse architectures. - **Original**: we lay down a principled basis for this framework and demonstrate its broad applicability across diverse neural architectures, including feedforward, convolutional, and transformer networks. - **Candidate**: theorem 7 let x and y be random elements of borel spaces x and y, respectively, and ga compact group acting measurably on x. assume that px is g-invariant, and pick a maximal invariant m : x →s, with sanother borel space. then py|x is g-invariant if and only if there exists a measurable function f :...

---

4. Monomial matrix group equivariant neural functional networks

URL: View paper

Prior Art Analysis

Monomial Matrix Functionals[75] demonstrates that prior work has already established formal theoretical foundations analyzing parameter spaces and characterizing symmetry groups in neural networks. The candidate paper explicitly characterizes maximal symmetry groups and their connection to functional equivalence for feedforward and convolutional networks, providing the same type of principled group-theoretic connection that the original paper claims as novel. Both papers analyze parameter spaces, define symmetry groups formally, introduce maximality concepts, and connect these to functional equivalence through group actions.

Evidence

Evidence 1 - **Rationale**: The candidate formally defines and characterizes symmetry groups preserved by activations, establishing the theoretical foundation for connecting symmetries to functional behavior. - **Original**: we examine the parameter space of a parameterized function, characterize its associated symmetry group, and introduce the formal notion of maximality within symmetry groups, establishing a direct connection to functional equivalence. - **Candidate**: definition 3.3. a matrix a ∈ gl(n) is said to be preserved by an activation σ if and only if σ(a · x) = a · σ(x) for all x ∈ rn. we adopt the term matrix group preserved by an activation from [68]. this term is then referred to as the intertwiner group of an activation in [24].

Evidence 2 - **Rationale**: Both papers introduce and formalize the notion of maximal symmetry groups with nearly identical definitions connecting to functional equivalence. - **Original**: definition 2.2(maximal symmetry group).a symmetry groupgis said to bemaximalif there exists a proper real algebraic varietyε⊊θsuch that, for allθ, ˜θ∈θ\ε, wheneverf(·;θ) =f(·; ˜θ), there existsg∈gwith ˜θ=gθ. - **Candidate**: remark 4.5 (maximality of g). the proof of proposition 4.4 can be found in appendix c.2. the group g defined above is even proved to be the maximal choice in the case: • σ = relu and nl ⩾ . . .⩾ n2 ⩾ n1 > n0 = 1 (see [12, 25]), or • σ = tanh (see [14, 22]). here, g is maximal in the sense that: if u...

5. Permutation Equivariant Neural Functionals

URL: View paper

Brief Assessment

[Final Audit Failure] The model insisted on a refutation claim but failed to provide verifiable evidence after multiple retries. Marked as cannot_refute for safety. Please manually verify the candidate text.

6. Complete Characterization of Gauge Symmetries in Transformer Architectures

URL: View paper

Prior Art Analysis

Transformer Gauge Symmetries[73] demonstrates that similar theoretical foundations connecting symmetry groups to functional equivalence were established prior to the original paper. Both papers analyze parameter spaces, characterize symmetry groups, and introduce notions of maximal symmetry groups to establish connections between group-theoretic symmetries and functional equivalence. The candidate paper provides a complete characterization of gauge symmetries in transformers through formal group-theoretic analysis, proving maximality of symmetry groups and their connection to functional equivalence, which overlaps substantially with the original paper's claimed contribution of establishing a principled theoretical foundation.

Evidence

Evidence 1 - **Rationale**: Both papers establish formal theoretical frameworks for analyzing symmetry groups in neural architectures, with the candidate specifically addressing transformers. - **Original**: we lay down a principled basis for this framework and demonstrate its broad applicability across diverse neural architectures, including feedforward, convolutional, and transformer networks - **Candidate**: we establish the complete gauge group structure for the canonical transformer family, which encompasses standard architectures including gpt-2, bert, llama, and qwen

Evidence 2 - **Rationale**: Both papers introduce and formalize the notion of maximal symmetry groups and establish their connection to functional equivalence, demonstrating prior work on this theoretical foundation. - **Original**: in section 2, we examine the parameter space of a parameterized function, characterize its associated symmetry group, and introduce the formal notion of maximality within symmetry groups, establishing a direct connection to functional equivalence - **Candidate**: we prove global maximality: the gauge group equals exactlygmax = ((gl(dk))h x (gl(dv))h) ⋊s h on the generic stratum where projection matrices have full column rank and head-wise attention controllability holds

Evidence 3 - **Rationale**: Both papers provide formal definitions of symmetry groups that preserve functional equivalence, showing similar theoretical approaches to characterizing parameter symmetries. - **Original**: definition 2.1(symmetry group).a groupgis called asymmetry groupof the functionfif gθ⊆[θ]for allθ∈θ. equivalently, for everyg∈gandθ∈θ, one hasf(·;gθ) =f(·;θ) - **Candidate**: definition 1 (standard gauge transformations)the standard gauge group gmax consists of transformations parametrized by(ai, ci) ∈gl (dk) xgl (dv)for each head i∈ {1, . . . , h}and permutationsσ∈sh

Evidence 4 - **Rationale**: Both papers formally define and prove maximality of symmetry groups, establishing that their characterized groups capture all functional equivalences. This demonstrates prior work on the theoretical foundation connecting maximal symmetry groups to functional equivalence. - **Original**: definition 2.2(maximal symmetry group).a symmetry groupgis said to bemaximalif there exists a proper real algebraic varietyε⊊θsuch that, for allθ, ˜θ∈θ\ε, wheneverf(·;θ) =f(·; ˜θ), there existsg∈gwith ˜θ=gθ - **Candidate**: theorem 2 (global maximality on the generic stratum)for the canonical transformer family satisfying assumptions a1-a4, a6-a8, the gauge group on the generic stratum θ0 equals exactlygmax = ((gl(dk))h x(gl(dv))h)⋊sh. no additional parameter symmetries exist beyond those in this group

7. Equivariant architectures for learning in deep weight spaces

URL: View paper

Prior Art Analysis

Deep Weight Spaces[70] demonstrates that prior work exists establishing the theoretical connection between symmetry groups and functional equivalence in neural networks. The candidate paper explicitly characterizes the symmetry group of weight spaces and provides a formal mathematical framework connecting these symmetries to functional equivalence, predating the original paper's contribution. Both papers analyze parameter spaces, characterize symmetry groups through permutation actions, and establish the connection to functional equivalence, with the candidate providing detailed mathematical formulations including group representations and equivariance properties.

Evidence

Evidence 1 - **Rationale**: Both papers identify the same fundamental challenge of capturing functional equivalence through symmetry structure, indicating Deep Weight Spaces[70] addressed this theoretical foundation earlier. - **Original**: a major challenge in designing metanetworks lies on how to capture functional equivalence - the fact that multiple distinct parameter configurations can realize the same input-output function - **Candidate**: the unique symmetry structure of deep weight spaces makes this design very challenging. if successful, such architectures would be capable of performing a wide range of intriguing tasks, from adapting a pre-trained network to a new domain to editing objects represented as functions (inrs or nerfs). ...

8. Equivariant matrix function neural networks

URL: View paper

Brief Assessment

Matrix Function Networks[74] focuses on equivariant matrix functions for graph neural networks, not on analyzing parameter spaces and symmetry groups that characterize functional equivalence in neural network weight spaces.

9. Adaptive knowledge assessment via symmetric hierarchical Bayesian neural networks with graph symmetry-aware concept dependencies

URL: View paper

Brief Assessment

Hierarchical Bayesian Symmetry[71] focuses on symmetric graph properties and symmetric concept representations in educational assessment systems, not on characterizing symmetry groups in neural network parameter spaces or establishing connections between group-theoretic symmetries and functional equivalence in neural architectures.

10. Symmetry in Neural Network Parameter Spaces

URL: View paper

Brief Assessment

[Final Audit Failure] The model insisted on a refutation claim but failed to provide verifiable evidence after multiple retries. Marked as cannot_refute for safety. Please manually verify the candidate text.

Contribution 3: General construction method for quasi-equivariant metanetworks

Description: The authors develop a practical construction framework for quasi-equivariant layers that can be applied to various neural architectures. The framework decomposes the design into group-valued maps and equivariant components, with concrete implementations for feedforward networks, CNNs, and transformers.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Neural Fields Meet Attention

URL: View paper

Brief Assessment

Neural Fields Attention[64] focuses on equivariance properties of transformer-based operators for learning neural fields (continuous functions), not on constructing quasi-equivariant metanetworks for processing pretrained network weights. The candidate addresses affine equivariance in field learning, while the original develops quasi-equivariance for weight-space metanetworks across feedforward, CNN, and transformer architectures.

2. Equivariant Neural Functional Networks for Transformers

URL: View paper

Brief Assessment

Equivariant Functional Transformers[63] focuses specifically on transformers and multihead attention architectures, not on a general construction framework applicable to feedforward networks, CNNs, and transformers as claimed in the original paper.

3. Long-Short-Range Fused Network for Heterogeneous Graph Representation Learning

URL: View paper

Brief Assessment

Long-Short-Range Fused[62] focuses on heterogeneous graph representation learning with aggregation layers and multi-head attention for graph structures, not on constructing quasi-equivariant metanetworks for neural network weight spaces.

4. Equivariant Diffusion-Based Sequential Hypergraph Neural Networks with Co-attention Fusion for Information Diffusion Prediction

URL: View paper

Brief Assessment

Diffusion Hypergraph Co-Attention[66] focuses on information diffusion prediction using hypergraph neural networks with attention mechanisms, not on constructing equivariant metanetworks for neural architecture weight spaces.

5. Graph Metanetworks for Processing Diverse Neural Architectures

URL: View paper

Brief Assessment

Graph Metanetworks Architectures[36] focuses on graph-based representations and standard GNN equivariance to graph automorphisms, not on quasi-equivariant frameworks that relax strict equivariance for metanetworks.

6. Permutation Equivariant Neural Functionals

URL: View paper

Brief Assessment

Permutation Equivariant Functionals[61] focuses on strict equivariance through parameter sharing schemes, not quasi-equivariance. The candidate constructs strictly equivariant nf-layers, whereas the original introduces quasi-equivariance to relax strict constraints.

7. GradMetaNet: An Equivariant Architecture for Learning on Gradients

URL: View paper

Brief Assessment

GradMetaNet[60] focuses on equivariant architectures for processing gradients of neural networks, not on constructing quasi-equivariant metanetworks for general weight-space learning. The candidate addresses gradient-space symmetries while the original addresses weight-space functional equivalence.

8. D3MES: Diffusion Transformer with multihead equivariant self-attention for 3D molecule generation

URL: View paper

Brief Assessment

D3MES[65] focuses on 3D molecule generation using diffusion models with multihead equivariant self-attention for molecular structures. It does not address metanetwork construction or quasi-equivariance frameworks for neural architectures like feedforward networks or transformers.

9. Dynamic Neural Graph: Facilitating Temporal Dynamics Learning in Deep Weight Space

URL: View paper

Brief Assessment

Dynamic Neural Graph[67] focuses on representing neural network parameters as dynamic graphs to capture temporal dynamics during inference, not on constructing quasi-equivariant metanetworks for various architectures. The candidate addresses graph-based processing of weight spaces rather than equivariance frameworks for metanetwork design.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Quasi-Equivariant Metanetworks View paper
- [1] Metapruning: Meta learning for automatic neural network channel pruning View paper
- [2] Equivariant deep weight space alignment View paper
- [3] Image-based plant disease identification by deep learning meta-architectures View paper
- [4] Hyper-Representations as Generative Models: Sampling Unseen Neural Network Weights View paper
- [5] Modularity in deep learning View paper
- [6] Meta-Sparsity: Learning Optimal Sparse Structures in Multi-task Networks through Meta-learning View paper
- [7] Scale equivariant graph metanetworks View paper
- [8] Improving Hyperparameter Optimization with Checkpointed Model Weights View paper
- [9] Rapid model architecture adaption for meta-learning View paper
- [10] Analysis of weighted networks View paper
- [11] Accelerating physics-informed neural network based 1D arc simulation by meta learning View paper
- [12] Weightnet: Revisiting the design space of weight networks View paper
- [13] On-NAS: On-device neural architecture search on memory-constrained intelligent embedded systems View paper
- [14] Meta Pruning via Graph Metanetworks : A Universal Meta Learning Framework for Network Pruning View paper
- [15] Meta Pruning via Graph Metanetworks: A Meta Learning Framework for Network Pruning View paper
- [16] Universal Neural Functionals View paper
- [17] Text2Weight: Bridging Natural Language and Neural Network Weight Spaces View paper
- [18] GL Equivariant Metanetworks for Learning on Low Rank Weight Spaces View paper
- [19] Network Dynamics-Based Framework for Understanding Deep Neural Networks View paper
- [20] Overlap-aware meta-learning attention to enhance hypergraph neural networks for node classification View paper
- [21] Exploring Local Transformation Shared Weights in Convolutional Neural Networks View paper
- [22] Generating Synaptic Weights from Task Specifications: A Computational Paradigm for Neural Network Initialization View paper
- [23] A deep patch network with spatiotemporal meta-parameter learning for soft sensor modeling of industrial processes View paper
- [24] Detection of food allergy using deep learning View paper
- [25] EMAS: Efficient Meta Architecture Search for Few-Shot Learning View paper
- [26] Hardware-adaptive efficient latency prediction for nas via meta-learning View paper
- [27] Help: Hardware-adaptive efficient latency prediction for nas via meta-learning View paper
- [28] Symmetry-Aware Graph Metanetwork Autoencoders: Model Merging through Parameter Canonicalization View paper
- [29] Learning to learn by jointly optimizing neural architecture and weights View paper
- [30] A neural network that embeds its own meta-levels View paper
- [31] FMW-Net: a first-order meta-weight-net approach for sample weighting View paper
- [32] Meta-learning approach to neural network optimization View paper
- [33] Metanetworks as Regulatory Operators: Learning to Edit for Requirement Compliance View paper
- [34] Training and Generating Neural Networks in Compressed Weight Space View paper
- [35] Symmetry-Aware Fully-Amortized Optimization with Scale Equivariant Graph Metanetworks View paper
- [36] Graph Metanetworks for Processing Diverse Neural Architectures View paper
- [37] Improving Recursive Transformers with Mixture of LoRAs View paper
- [38] Non-redundant genetic coding of neural networks View paper
- [39] Element-Wise Multiplicative Operators in Vision, Language, and Multimodal Learning View paper
- [40] On the geometry of feedforward neural network error surfaces View paper
- [41] Neural Functional Transformers View paper
- [42] Deep E-Learning RecommendNet: An Acute E-Learning Recommendation System with Meta-Heuristic-Based Hybrid Deep Learning Architecture View paper
- [43] Generating Neural Networks with Neural Networks View paper
- [44] A connectionist approach to speech recognition View paper
- [45] Engineering artificial systems with natural intelligence View paper
- [46] General-Purpose In-Context Learning by Meta-Learning Transformers View paper
- [47] Transformations of sigmaâ□□pi nets: obtaining reflected functions by reflecting weight matrices View paper
- [48] Metanetwork: A novel approach to interpreting ANNs View paper
- [49] Efficient Architecture Search by Network Transformation View paper
- [50] Relaxing equivariance constraints with non-stationary continuous filters View paper
- [51] AV-NAS: Audio-Visual Multi-Level Semantic Neural Architecture Search for Video Hashing View paper
- [52] Symmetry breaking and equivariant neural networks View paper
- [53] Approximately equivariant graph networks View paper
- [54] Equivariance-aware architectural optimization of neural networks View paper
- [55] On the scale invariance in state of the art CNNs trained on ImageNet View paper
- [56] Self-supervised image denoising with downsampled invariance loss and conditional blind-spot network View paper
- [57] Sharp minima can generalize for deep nets View paper
- [58] Weakly connected neural networks View paper
- [59] Relaxed Equivariant Graph Neural Networks View paper
- [60] GradMetaNet: An Equivariant Architecture for Learning on Gradients View paper
- [61] Permutation Equivariant Neural Functionals View paper
- [62] Long-Short-Range Fused Network for Heterogeneous Graph Representation Learning View paper
- [63] Equivariant Neural Functional Networks for Transformers View paper
- [64] Neural Fields Meet Attention View paper
- [65] D3MES: Diffusion Transformer with multihead equivariant self-attention for 3D molecule generation View paper
- [66] Equivariant Diffusion-Based Sequential Hypergraph Neural Networks with Co-attention Fusion for Information Diffusion Prediction View paper
- [67] Dynamic Neural Graph: Facilitating Temporal Dynamics Learning in Deep Weight Space View paper

- [68] Symmetry in Neural Network Parameter Spaces View paper
- [69] Universal approximations of invariant maps by neural networks View paper
- [70] Equivariant architectures for learning in deep weight spaces View paper
- [71] Adaptive knowledge assessment via symmetric hierarchical Bayesian neural networks with graph symmetry-aware concept dependencies View paper
- [72] Probabilistic symmetries and invariant neural networks View paper
- [73] Complete Characterization of Gauge Symmetries in Transformer Architectures View paper
- [74] Equivariant matrix function neural networks View paper
- [75] Monomial matrix group equivariant neural functional networks View paper
- [76] A symmetry-aware exploration of bayesian neural network posteriors View paper