Novelty Assessment Report

Paper: LipNeXt: Scaling up Lipschitz-based Certified Robustness to Billion-parameter Models
PDF URL: https://openreview.net/pdf?id=2fgzEJPH02
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-29

Abstract

Lipschitz-based certification offers efficient, deterministic robustness guarantees but has struggled to scale in model size, training efficiency, and ImageNet performance. We introduce \emph{LipNeXt}, the first \emph{constraint-free} and \emph{convolution-free} 1-Lipschitz architecture for certified robustness. LipNeXt is built using two techniques: (1) a manifold optimization procedure that updates parameters directly on the orthogonal manifold and (2) a \emph{Spatial Shift Module} to model spatial pattern without convolutions. The full network uses orthogonal projections, spatial shifts, a simple 1-Lipschitz β-Abs nonlinearity, and L_2 spatial pooling to maintain tight Lipschitz control while enabling expressive feature mixing. Across CIFAR-10/100 and Tiny-ImageNet, LipNeXt achieves state-of-the-art clean and certified robust accuracy (CRA), and on ImageNet it scales to 1–2B large models, improving CRA over prior Lipschitz models (e.g., up to $+8\%$ at $\varepsilon{=}1$) while retaining efficient, stable low-precision training. These results demonstrate that Lipschitz-based certification can benefit from modern scaling trends without sacrificing determinism or efficiency.

Core Task Landscape

This paper addresses: **Lipschitz-based Certified Robustness for Neural Networks**

A total of **50 papers** were analyzed and organized into a taxonomy with **19 categories**.

Taxonomy Overview

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

- **Lipschitz Constant Estimation and Computation**
- **Training Methods with Lipschitz Regularization or Constraints**
- **Lipschitz-constrained Neural Network Architectures**
- **Domain-specific and Extended Applications**
- **Surveys and Position Papers**

Complete Taxonomy Tree

- Lipschitz-based Certified Robustness for Neural Networks Survey Taxonomy
- Lipschitz Constant Estimation and Computation
  - Exact and Tight Bound Computation (4 papers)
  - [9] Efficient and accurate estimation of lipschitz constants for deep neural networks (Fazlyab, 2019) View paper
  - [16] ECLipsE: Efficient Compositional Lipschitz Constant Estimation for Deep Neural Networks (S. Sivaranjani, 2024) View paper
  - [22] Computable Lipschitz Bounds for Deep Neural Networks (Pintore, 2024) View paper
  - [45] MIQCQP reformulation of the ReLU neural networks Lipschitz constant estimation problem (Sbihi, 2024) View paper
  - Sampling-based and Local Estimation (2 papers)
  - [8] Evaluating the Robustness of Neural Networks: An Extreme Value Theory Approach (Weng, 2018) View paper
  - [50] Efficiently Computing Local Lipschitz Constants of Neural Networks via Bound Propagation (Shi, 2022) View paper
  - Architecture-specific and Control-theoretic Estimation (5 papers)
  - [20] Estimating neural network robustness via lipschitz constant and architecture sensitivity (Abulikemu Abuduweili, 2024) View paper
  - [40] Lipschitz-Based Robustness Certification for Recurrent Neural Networks via Convex Relaxation (Schiffer Johannes, 2025) View paper
  - [46] Validated Computation of Lipschitz Constant of Recurrent Neural Networks (Yuhua Guo, 2023) View paper
  - [47] Lipschitz constant estimation for general neural network architectures using control tools (Pauli, 2024) View paper
  - [49] Lipschitz Constants of Hybrid Zonotope Representations of Feedforward Neural Networks (Justin Chen, 2025) View paper
  - Theoretical Analysis of Lipschitz Properties (3 papers)
  - [11] A law of robustness for two-layers neural networks (Bubeck, 2021) View paper
  - [14] Lipschitz regularity of deep neural networks: analysis and efficient estimation (Scaman, 2018) View paper
  - [23] On the robustness of a neural network (El-Mhamdi, 2017) View paper
- Training Methods with Lipschitz Regularization or Constraints
  - Margin-based and Calibrated Loss Training (3 papers)
  - [3] Certified robust models with slack control and large Lipschitz constants (Losch, 2023) View paper
  - [6] Lipschitz-margin training: Scalable certification of perturbation invariance for deep neural networks (Yusuke Tsuzuku, 2018) View paper
  - [12] Certified Robustness via Dynamic Margin Maximization and Improved Lipschitz Regularization (Fazlyab, 2023) View paper
  - Constrained Optimization and Semidefinite Programming (2 papers)
  - [1] Training robust neural networks using Lipschitz bounds (Patricia Pauli, 2021) View paper
  - [21] Neural network training under semidefinite constraints (Patricia Pauli, 2022) View paper

- Local Lipschitz Bounds and Adaptive Training (1 papers)
- [17] Training Certifiably Robust Neural Networks with Efficient Local Lipschitz Bounds (Huang, 2021) View paper
- General Robustness Training via Lipschitz Control (3 papers)
- [13] Towards robust neural networks with lipschitz continuity (Muhammad Usama, 2018) View paper
- [18] Towards improving robustness of deep neural networks to adversarial perturbations (Sajjad Amini, 2020) View paper
- [38] Data-driven lipschitz continuity: A cost-effective approach to improve adversarial robustness (Chen, 2024) View paper

- Lipschitz-constrained Neural Network Architectures
  - Orthogonal and Cayley-based Parameterizations ★ (2 papers)
  - [0] LipNeXt: Scaling up Lipschitz-based Certified Robustness to Billion-parameter Models (Anon et al., 2026) View paper
  - [30] Lipschitz-Bounded 1D Convolutional Neural Networks using the Cayley Transform and the Controllability Gramian (Patricia Pauli, 2023) View paper
  - 1-Lipschitz Network Designs and Activation Functions (4 papers)
  - [4] Globally-robust neural networks (Leino, 2021) View paper
  - [33] Pay attention to your loss: understanding misconceptions about lipschitz neural networks (Louis Bethune, 2022) View paper
  - [39] 1-lipschitz layers compared: Memory speed and certifiable robustness (Bernd Prach, 2024) View paper
  - [44] Universal Lipschitz Approximation in Bounded Depth Neural Networks (Huster, 2019) View paper
  - Randomized Smoothing with Lipschitz Constraints (1 papers)
  - [19] SPLITZ: Certifiable Robustness via Split Lipschitz Randomized Smoothing (Meiyu Zhong, 2025) View paper
  - Pre-trained and Large-scale Lipschitz Models (2 papers)
  - [24] Certvit: Certified robustness of pre-trained vision transformers (Gupta, 2023) View paper
  - [27] A recipe for improved certifiable robustness (Hu Kai, 2023) View paper
- Domain-specific and Extended Applications
  - Recurrent and Sequential Models (3 papers)
  - [15] Lipsnet: a smooth and robust neural network with adaptive lipschitz constant for high accuracy optimal control (X Song, 2023) View paper
  - [32] Improve Robustness of Reinforcement Learning against Observation Perturbations via lâ Lipschitz Policy Networks (Buqing Nie, 2023) View paper
  - [41] Adaptive Neural Stochastic Control With Lipschitz Constant Optimization (Lian Geng, 2024) View paper
  - Graph Neural Networks and Non-Euclidean Domains (2 papers)
  - [28] Improving Robustness of Hyperbolic Neural Networks by Lipschitz Analysis (Yuekang Li, 2024) View paper
  - [35] Robust graph neural networks via probabilistic lipschitz constraints (Raghu Arghal, 2022) View paper
  - Specialized Architectures and Quantization (2 papers)
  - [5] Lipschitz continuity retained binary neural network (Shang, 2022) View paper
  - [25] Exploiting connections between Lipschitz structures for certifiably robust deep equilibrium models (A Havens, 2023) View paper
  - Ensemble Methods and Generalization (2 papers)
  - [29] Certifying ensembles: A general certification theory with S-Lipschitzness (Petrov, 2023) View paper
  - [48] The robust way to stack and bag: the local Lipschitz way (Tholeti, 2022) View paper
  - Domain-independent Certification and Cross-domain Robustness (3 papers)
  - [7] Verification of Geometric Robustness of Neural Networks via Piecewise Linear Approximation and Lipschitz Optimisation (Zheng Yang, 2024) View paper
  - [10] Achieving domain-independent certified robustness via knowledge continuity (Alan Sun, 2024) View paper
  - [43] Certifying Robustness in NLP Classifiers via Lipschitz Constraints (Zamora, 2025) View paper
  - Application-specific Implementations (5 papers)
  - [26] Applications of Lipschitz neural networks to the Run 3 LHCb trigger system (B. Delaney, 2023) View paper
  - [34] Lipschitz-based robustness estimation for hyperdimensional learning. (Calvin Yeung, 2025) View paper
  - [36] EMG-Based Automatic Gesture Recognition Using Lipschitz-Regularized Neural Networks (Ana NeacÅu, 2023) View paper
  - [37] Robust One-Class Classification with Signed Distance Function using 1-Lipschitz Neural Networks (Bethune, 2023) View paper
  - [42] Towards Certifiably Robust Face Recognition (Seunghun Paik, 2024) View paper
- Surveys and Position Papers (2 papers)
  - [2] Adversarial robustness of neural networks from the perspective of lipschitz calculus: A survey (Monty-Maximilian ZÃ¼hlke, 2025) View paper
  - [31] Is Certifying Robustness Still Worthwhile? (R Mangal, 2023) View paper

Narrative

Core task: Lipschitz-based certified robustness for neural networks. The field centers on bounding the Lipschitz constant of neural networks to provide formal guarantees against adversarial perturbations. The taxonomy reveals several complementary research directions: one branch focuses on estimating and computing Lipschitz constants efficiently (e.g., Efficient Lipschitz Estimation[9], Computable Lipschitz Bounds[22]), another explores training methods that incorporate Lipschitz regularization or constraints to improve robustness during optimization (e.g., Lipschitz Margin Training[6], Slack Control Lipschitz[3]), and a third develops specialized architectures that enforce Lipschitz constraints by design through orthogonal or Cayley-based parameterizations (e.g., Cayley Transform Convolutions[30]). Additional branches address domain-specific applications—ranging from face recognition (Certifiable Face Recognition[42]) to NLP (NLP Lipschitz Certification[43])—and provide broader surveys (Lipschitz Robustness Survey[2]) that synthesize these threads.

A particularly active line of work explores the trade-offs between tight Lipschitz bounds and model expressiveness. Some studies pursue globally robust architectures with strict 1-Lipschitz layers (Globally Robust Networks[4], 1-Lipschitz Layers Compared[39]), while others investigate local or adaptive bounds that balance certification strength with practical accuracy (Local Lipschitz Bounds[17], Dynamic Margin Maximization[12]). Within this landscape, LipNeXt[0] sits in the architectures branch alongside Cayley Transform Convolutions[30], emphasizing orthogonal and Cayley-based parameterizations to enforce Lipschitz constraints structurally. Compared to training-centric approaches like Slack Control Lipschitz[3] or estimation-focused methods like Efficient Lipschitz Estimation[9], LipNeXt[0] prioritizes architectural design to achieve certified robustness, reflecting a growing interest in building guarantees directly into network layers rather than relying solely on post-hoc verification or regularization penalties.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Lipschitz-Bounded 1D Convolutional Neural Networks using the Cayley Transform and the Controllability Gramian

Authors: Patricia Pauli, Ruigang Wang, Ian R. Manchester, Frank AllgÃ¶wer, I. Manchester, et al. (6 authors total) | **Year/Venue**: 2023 • IEEE Conference on Decision and Control | **URL**: View paper

Abstract

We establish a layer-wise parameterization for 1D convolutional neural networks (CNNs) with built-in end-to-end robustness guarantees. In doing so, we use the Lipschitz constant of the input-output mapping characterized by a CNN as a robustness measure. We base our parameterization on the Cayley transform that parameterizes orthogonal matrices and the controllability Gramian of the state space representation of the convolutional layers. The proposed parameterization by design fulfills linear mat...

Relationship Analysis

Both papers belong to the Orthogonal and Cayley-based Parameterizations category, using orthogonal weight matrices to maintain Lipschitz bounds. The candidate paper focuses on 1D convolutional neural networks using the Cayley transform combined with controllability Gramians for state-space representations of convolutional layers, while the original paper (LipNeXt) proposes a constraint-free manifold optimization approach with a convolution-free Spatial Shift Module, scaling to billion-parameter models on ImageNet. The key difference is that the candidate uses explicit Cayley-based re-parameterization for 1D CNNs with state-space formulations, whereas the original employs direct manifold optimization with spatial shifts to achieve scalability to much larger models and datasets.

Contributions Analysis

Overall novelty summary. The paper introduces LipNeXt, a constraint-free and convolution-free 1-Lipschitz architecture combining manifold optimization for orthogonal parameters with a novel Spatial Shift Module. Within the taxonomy, it resides in the 'Orthogonal and Cayley-based Parameterizations' leaf under 'Lipschitz-constrained Neural Network Architectures'. This leaf contains only two papers total, indicating a relatively sparse research direction focused specifically on orthogonal weight parameterizations for Lipschitz control. The sibling work explores Cayley transforms, suggesting the area is emerging but not yet crowded.

The taxonomy reveals that LipNeXt sits within a broader architectures branch containing four subcategories: orthogonal parameterizations, 1-Lipschitz network designs (four papers), randomized smoothing hybrids (one paper), and pre-trained large-scale models (two papers). Neighboring branches include training methods with Lipschitz regularization (seven papers across four leaves) and estimation techniques (eleven papers across four leaves). The scope notes clarify that LipNeXt's orthogonal parameterization distinguishes it from general 1-Lipschitz designs that may use other layer compositions, and from training methods that add regularization to standard architectures rather than building constraints into the structure.

Among twenty candidates examined, the manifold optimization contribution shows overlap with two prior works, while the Spatial Shift Module was not evaluated against any candidates (zero examined). The architecture-level contribution (achieving state-of-the-art certified robustness at scale) was assessed against ten candidates with no clear refutations found. This suggests that among the limited semantic matches retrieved, the core architectural innovation and scaling results appear less directly anticipated, though the orthogonal parameterization technique itself has documented precedents. The analysis explicitly covers top-K semantic search plus citation expansion, not an exhaustive literature review.

Given the sparse taxonomy leaf (two papers) and limited search scope (twenty candidates), the work appears to occupy a relatively underexplored intersection of orthogonal parameterizations and large-scale certified robustness. The Spatial Shift Module and scaling achievements show no clear prior overlap within the examined set, though the manifold optimization approach has established antecedents. A broader search might reveal additional related work in computer vision or efficient architectures not captured by Lipschitz-focused queries.

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

Contribution 1: Constraint-free manifold optimization for orthogonal parameters

Description: The authors propose a manifold optimization procedure that updates orthogonal parameters directly on the orthogonal manifold, avoiding re-parameterization constraints. They introduce FastExp, a norm-adaptive Taylor series approximation of the matrix exponential, combined with periodic polar retraction and manifold-adapted Lookahead stabilization to enable efficient and stable training of large-scale 1-Lipschitz 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. Online optimization over Riemannian manifolds

URL: View paper

Brief Assessment

Riemannian Online Optimization[53] focuses on online optimization problems over Riemannian manifolds with regret analysis, not on training large-scale neural networks with orthogonal constraints for certified robustness. The candidate addresses different optimization settings (online convex optimization) rather than the offline training framework with FastExp approximations and stabilization techniques proposed in the original paper.

---

2. Cross-Coupling Matrix Reconfiguration Using the Levenbergâ Marquardt Algorithm on Orthogonal Groups

URL: View paper

Brief Assessment

Levenberg Marquardt Orthogonal[54] applies manifold optimization to coupling matrix reconfiguration in filter design, not to neural network training or certified robustness. The application domain and objectives are fundamentally different from the original paper's focus on scaling Lipschitz-based certified robustness.

---

3. Amortized eigendecomposition for neural networks

URL: View paper

Brief Assessment

Amortized Eigendecomposition[55] focuses on eigendecomposition within neural networks using QR decomposition for reparameterization, not on manifold optimization for orthogonal parameters using matrix exponential approximations as in the original paper.

---

4. Cheap orthogonal constraints in neural networks: A simple parametrization of the orthogonal and unitary group

URL: View paper

Prior Art Analysis

Cheap Orthogonal Constraints[56] demonstrates that manifold optimization for orthogonal parameters using matrix exponential approximations was already proposed and implemented prior to the original paper. The candidate paper presents a parametrization approach that updates orthogonal parameters directly on the orthogonal manifold using the exponential map, combined with efficient approximations via Padé approximants and the scale-squaring trick. This directly addresses the same core problem of constraint-free optimization on the orthogonal manifold that the original paper claims as novel. Both papers use the matrix exponential to transform constrained optimization into unconstrained optimization, and both employ approximation techniques to make the exponential computation tractable.

Evidence

Evidence 1 - **Rationale**: Both papers propose the same fundamental approach: using the exponential map to transform constrained optimization on the orthogonal manifold into unconstrained optimization in Euclidean space. The candidate paper explicitly presents this as their core contribution in 2019, predating the original paper's claim of novelty. - **Original**: we propose directly optimizing orthogonal matrices on the orthogonal manifold. although orthogonal manifold optimization is a mature technique, to our knowledge it has not been exploited for certification. - **Candidate**: we introduce a novel approach to perform firstorder optimization with orthogonal and unitary constraints. this approach is based on a parametrization stemming from lie group theory through the exponential map. the parametrization transforms the constrained optimization problem into an unconstrained o...

Evidence 2 - **Rationale**: The candidate paper already describes efficient approximation methods for the matrix exponential using Padé approximants and the scale-squaring trick, which are the same class of techniques the original paper uses for efficient computation. - **Original**: we further observe that in the large-model regime, where learning rates are small, the matrix exponential can be accurately and efficiently approximated. combining these ideas, we reduce the additional per-update overhead to at most five matrix multiplications. - **Candidate**: pad´e approximants are rational approximations of the form exp(a) ≈ pn(a)qn(a)-1 for polynomials pn,qn of degree n. a pad ´e approximant of degree nagrees with the taylor expansion of the exponential to degree 2n. the cayley transform is the pad´e approximant of degree 1. these methods and their impl...

Evidence 3 - **Rationale**: The candidate paper provides the mathematical foundation for parametrizing the orthogonal group using manifold optimization, which is the theoretical basis for the original paper's claimed contribution. - **Original**: a manifold optimization procedure that updates parameters directly on the orthogonal manifold and (2) aspatial shift moduleto model spatial pattern without convolutions. the full network uses orthogonal projections, spatial shifts, a simple 1-lipschitzβ-abs nonlinearity, andl 2 spatial pooling to ma... - **Candidate**: we are interested in the study of parametrizations of the special orthogonal group so(n) = {b ∈rnxn |b⊺b = i, det(b) = 1} and the unitary group u(n) = {b ∈cnxn |b∗b = i}. these two sets are compact and connected lie groups. furthermore, when seen as submanifolds ofrnxn(resp. cnxn) equipped with the ...

Evidence 4 - **Rationale**: The candidate paper explicitly addresses the same problem of avoiding hard orthogonality constraints through their parametrization approach, which is a key claimed advantage in the original paper. - **Original**: examining existing work, we find that orthogonal matrices are fundamental to building 1-lipschitz networks because they enable tight lipschitz bounds. however, they are also a major bottleneck that prevents lipschitz-based certification from scaling. existing methods either explicitly re-parameteriz... - **Candidate**: the parametrization does not create additional minima or saddle points in the main parametrization region. 3. it is possible to use a structured initializer to take advantage of the structure of the eigenvalues of the orthogonal matrix. 4. other approaches need to enforce hard orthogonality constrai...

---

5. Smoothly Evolving Geodesics in the Special Orthogonal Group: Definitions, Computations and Applications

URL: View paper

Brief Assessment

Geodesics Special Orthogonal[58] focuses on theoretical foundations of geodesics and matrix exponential/logarithm computations on SO(n), not on neural network training or certified robustness applications.

---

6. Fast and accurate optimization on the orthogonal manifold without retraction

URL: View paper

Prior Art Analysis

Orthogonal Manifold Optimization[51] demonstrates that constraint-free manifold optimization for orthogonal parameters was previously proposed. The candidate paper presents the 'landing algorithm' which updates orthogonal parameters directly on the manifold without using retractions, avoiding expensive operations like matrix exponentials. This directly challenges the novelty claim of the original paper's manifold optimization procedure. Both papers address the same core problem: updating orthogonal parameters on the manifold efficiently without re-parameterization constraints. The candidate's landing algorithm uses a potential energy-based approach with matrix multiplications, while the original's FastExp uses Taylor series approximations, but both fundamentally propose constraint-free optimization on the orthogonal manifold.

Evidence

Evidence 1 - **Rationale**: Both papers propose constraint-free approaches to orthogonal manifold optimization. The candidate explicitly introduces a method that avoids retractions and expensive operations, which is the same core contribution claimed by the original paper. - **Original**: we proposelipnext, the first convolution-free and constraint-free architecture for certified robustness. benefiting from this design, we scale to billionparameter models and observe non-saturating gains with increasing model size. - **Candidate**: we consider the problem of minimizing a function over the manifold of orthogonal matrices. the majority of algorithms for this problem compute a direction in the tangent space, and then use a retraction to move in that direction while staying on the manifold. unfortunately, the numerical computation...

Evidence 2 - **Rationale**: Both papers explicitly contrast their constraint-free approaches against traditional constrained/retraction-based methods, demonstrating that the candidate paper already established this distinction. - **Original**: as discussed in section 2.2, prior works use re-parameterization to learn orthogonal matrices explicitly or implicitly. we name these asconstrainedapproaches, as the optimization variables do not reside directly on the orthogonal manifoldm d. in contrast, we adopt aconstraint-freemanifold optimizati... - **Candidate**: all these methods arefeasible, i.e. generate a sequence of iteratesxk where each iterate is inop. unlike what we assume in the first sentence of the present article, they do not need the functionf to be defined outside op. this comes with a computational drawback: in order to computexk+1 from xk, one ...

Evidence 3 - **Rationale**: Both papers address the computational bottleneck of expensive operations in manifold optimization. The candidate's landing algorithm uses only matrix multiplications, while the original uses Taylor approximations, but both claim to solve the same efficiency problem in constraint-free manifold optimization. - **Original**: a significant bottleneck in manifold optimization is the computational cost of the matrix exponential,exp(·)in equation 3. section 2.2 discussed solutions from prior work to address this issue. however, we observe that there exists a very simple solution in the context of training neural networks. a... - **Candidate**: we stress that the fieldλ is not the riemannian gradient nor the euclidean gradient of a function (its jacobian is not symmetric). in particular, the landing flow doesnot have the same trajectory as the euclidean gradient flow associated to the functionf(x)+ λn(x). before we move on to the analysis of ...

---

7. Numerical Approaches for Constrained and Unconstrained, Static Optimization on the Special Euclidean Group SE (3)

URL: View paper

Brief Assessment

Special Euclidean Optimization[59] focuses on numerical optimization methods for the SE(3) group (rigid body transformations), not on training neural networks with orthogonal weight matrices for certified robustness.

---

8. Quotient Geometry of Bounded or Fixed-Rank Correlation Matrices

URL: View paper

Brief Assessment

Correlation Matrix Geometry[57] focuses on quotient geometry of correlation matrices and Riemannian optimization for computing distances and Frechet means in orbit spaces. It does not address constraint-free optimization of orthogonal parameters in neural networks or matrix exponential approximations for training 1-Lipschitz networks.

---

9. A feasible method for optimization with orthogonality constraints

URL: View paper

Brief Assessment

Orthogonality Constraints Method[52] focuses on optimization with orthogonality constraints using Cayley transform and curvilinear search, but does not address the specific techniques in the original paper: FastExp (norm-adaptive Taylor series approximation), periodic polar retraction, or manifold-adapted Lookahead stabilization for training large-scale 1-Lipschitz neural networks.

---

10. A randomized feasible algorithm for optimization with orthogonal constraints

URL: View paper

Brief Assessment

Randomized Orthogonal Algorithm[60] focuses on a randomized feasible algorithm for optimization with orthogonal constraints, while the original paper proposes a specific manifold optimization procedure with FastExp approximation, polar retraction, and Lookahead stabilization for training 1-Lipschitz networks. The candidate's minimal context does not provide sufficient detail to assess whether it addresses the same technical contributions.

Contribution 2: Spatial Shift Module for convolution-free spatial mixing

Description: The authors design a parameter-free spatial mixing operator based on circular shifts applied to partitioned feature channels. They provide theoretical justification via Theorem 1, showing that norm-preserving depthwise convolutions reduce to spatial shifts, and combine this module with positional encoding to model spatial patterns without convolutions while maintaining tight 1-Lipschitz bounds.

This contribution was assessed against **0 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

Contribution 3: LipNeXt architecture achieving state-of-the-art certified robustness at scale

Description: The authors introduce LipNeXt, the first constraint-free and convolution-free 1-Lipschitz architecture for certified robustness. By integrating manifold optimization and the Spatial Shift Module with orthogonal projections and beta-Abs nonlinearity, LipNeXt achieves state-of-the-art certified robust accuracy and clean accuracy across CIFAR-10/100, Tiny-ImageNet, and ImageNet, successfully scaling to 1-2 billion parameters with efficient low-precision training.

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. Achieving domain-independent certified robustness via knowledge continuity

URL: View paper

Brief Assessment

Knowledge Continuity Robustness[10] proposes a fundamentally different approach to certified robustness based on knowledge continuity rather than Lipschitz constraints. While both address certified robustness, Knowledge Continuity Robustness[10] focuses on domain-independent guarantees through loss function variability in representation spaces, not on scaling Lipschitz-constrained architectures to billion parameters with specific architectural innovations like spatial shift modules and manifold optimization.

---

2. Training robust neural networks using Lipschitz bounds

URL: View paper

Brief Assessment

Training Lipschitz Bounds[1] focuses on training procedures using ADMM with SDP-based Lipschitz regularization for small-scale networks ($\leq$50 neurons, MNIST with pooling). It does not propose large-scale architectures, constraint-free manifold optimization, or convolution-free designs like LipNeXt's spatial shift module and billion-parameter scaling.

---

3. Globally-robust neural networks

URL: View paper

Brief Assessment

Globally Robust Networks[4] focuses on a different architectural approach using global Lipschitz bounds with a special $\perp$ class for certification, rather than the constraint-free manifold optimization and spatial shift modules that define LipNeXt's novelty.

---

4. A closer look at accuracy vs. robustness

URL: View paper

Brief Assessment

Accuracy versus Robustness[62] focuses on the theoretical relationship between accuracy and robustness through local Lipschitzness and data separation properties, without proposing specific architectures. The candidate does not present a Lipschitz-constrained architecture design for certified robustness at scale.

---

5. Certified robust models with slack control and large Lipschitz constants

URL: View paper

Brief Assessment

Slack Control Lipschitz[3] focuses on calibrated loss functions and slack control for Lipschitz-constrained models, not on architectural innovations like constraint-free manifold optimization or convolution-free spatial shift modules that enable billion-parameter scaling.

6. Adversarial robustness of neural networks from the perspective of lipschitz calculus: A survey

URL: View paper

Brief Assessment

Lipschitz Robustness Survey[2] is a survey paper that reviews existing techniques for Lipschitz-based robustness, including estimation algorithms, regularization techniques, and robustness guarantees. It does not present a novel architecture or experimental results that would challenge the novelty of LipNeXt's specific architectural innovations (constraint-free manifold optimization, spatial shift module, billion-parameter scaling).

7. Robust and provably monotonic networks

URL: View paper

Brief Assessment

Monotonic Networks[61] focuses on weight normalization for Lipschitz control in dense networks with monotonicity constraints for particle physics applications, not on scaling certified robustness architectures to billion parameters with convolution-free designs and manifold optimization.

8. SPLITZ: Certifiable Robustness via Split Lipschitz Randomized Smoothing

URL: View paper

Brief Assessment

SPLITZ[19] focuses on combining Lipschitz-constrained training with randomized smoothing for certified robustness, rather than proposing a novel Lipschitz-constrained architecture. The candidate does not present an alternative 1-Lipschitz architecture design that would challenge LipNeXt's novelty claims about constraint-free, convolution-free architectures with manifold optimization and spatial shift modules.

9. Unlocking deterministic robustness certification on imagenet

URL: View paper

Brief Assessment

ImageNet Robustness Certification[63] focuses on LiResNet architecture with linear residual blocks and EMMA loss for scaling deterministic certification to ImageNet. LipNeXt introduces a fundamentally different approach using constraint-free manifold optimization and spatial shift modules without convolutions, representing distinct technical innovations rather than prior work that would refute LipNeXt's novelty claims.

10. Estimating neural network robustness via lipschitz constant and architecture sensitivity

URL: View paper

Brief Assessment

Architecture Sensitivity Estimation[20] focuses on analytical estimation of Lipschitz constants for robustness analysis in MLPs, not on designing novel 1-Lipschitz architectures with manifold optimization and spatial shift modules for certified robustness at billion-parameter scale.

Appendix: Text Similarity Detection

Textual similarity detection checked 21 papers and found 1 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. Fast and accurate optimization on the orthogonal manifold without retraction

Detected in: Contribution: contribution_1

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] LipNeXt: Scaling up Lipschitz-based Certified Robustness to Billion-parameter Models View paper
- [1] Training robust neural networks using Lipschitz bounds View paper
- [2] Adversarial robustness of neural networks from the perspective of lipschitz calculus: A survey View paper
- [3] Certified robust models with slack control and large Lipschitz constants View paper
- [4] Globally-robust neural networks View paper
- [5] Lipschitz continuity retained binary neural network View paper
- [6] Lipschitz-margin training: Scalable certification of perturbation invariance for deep neural networks View paper
- [7] Verification of Geometric Robustness of Neural Networks via Piecewise Linear Approximation and Lipschitz Optimisation View paper
- [8] Evaluating the Robustness of Neural Networks: An Extreme Value Theory Approach View paper
- [9] Efficient and accurate estimation of lipschitz constants for deep neural networks View paper
- [10] Achieving domain-independent certified robustness via knowledge continuity View paper
- [11] A law of robustness for two-layers neural networks View paper
- [12] Certified Robustness via Dynamic Margin Maximization and Improved Lipschitz Regularization View paper
- [13] Towards robust neural networks with lipschitz continuity View paper
- [14] Lipschitz regularity of deep neural networks: analysis and efficient estimation View paper
- [15] Lipsnet: a smooth and robust neural network with adaptive lipschitz constant for high accuracy optimal control View paper
- [16] ECLipsE: Efficient Compositional Lipschitz Constant Estimation for Deep Neural Networks View paper
- [17] Training Certifiably Robust Neural Networks with Efficient Local Lipschitz Bounds View paper
- [18] Towards improving robustness of deep neural networks to adversarial perturbations View paper
- [19] SPLITZ: Certifiable Robustness via Split Lipschitz Randomized Smoothing View paper
- [20] Estimating neural network robustness via lipschitz constant and architecture sensitivity View paper

- [21] Neural network training under semidefinite constraints View paper
- [22] Computable Lipschitz Bounds for Deep Neural Networks View paper
- [23] On the robustness of a neural network View paper
- [24] Certvit: Certified robustness of pre-trained vision transformers View paper
- [25] Exploiting connections between Lipschitz structures for certifiably robust deep equilibrium models View paper
- [26] Applications of Lipschitz neural networks to the Run 3 LHCb trigger system View paper
- [27] A recipe for improved certifiable robustness View paper
- [28] Improving Robustness of Hyperbolic Neural Networks by Lipschitz Analysis View paper
- [29] Certifying ensembles: A general certification theory with S-Lipschitzness View paper
- [30] Lipschitz-Bounded 1D Convolutional Neural Networks using the Cayley Transform and the Controllability Gramian View paper
- [31] Is Certifying Robustness Still Worthwhile? View paper
- [32] Improve Robustness of Reinforcement Learning against Observation Perturbations via lâ Lipschitz Policy Networks View paper
- [33] Pay attention to your loss: understanding misconceptions about lipschitz neural networks View paper
- [34] Lipschitz-based robustness estimation for hyperdimensional learning. View paper
- [35] Robust graph neural networks via probabilistic lipschitz constraints View paper
- [36] EMG-Based Automatic Gesture Recognition Using Lipschitz-Regularized Neural Networks View paper
- [37] Robust One-Class Classification with Signed Distance Function using 1-Lipschitz Neural Networks View paper
- [38] Data-driven lipschitz continuity: A cost-effective approach to improve adversarial robustness View paper
- [39] 1-lipschitz layers compared: Memory speed and certifiable robustness View paper
- [40] Lipschitz-Based Robustness Certification for Recurrent Neural Networks via Convex Relaxation View paper
- [41] Adaptive Neural Stochastic Control With Lipschitz Constant Optimization View paper
- [42] Towards Certifiably Robust Face Recognition View paper
- [43] Certifying Robustness in NLP Classifiers via Lipschitz Constraints View paper
- [44] Universal Lipschitz Approximation in Bounded Depth Neural Networks View paper
- [45] MIQCQP reformulation of the ReLU neural networks Lipschitz constant estimation problem View paper
- [46] Validated Computation of Lipschitz Constant of Recurrent Neural Networks View paper
- [47] Lipschitz constant estimation for general neural network architectures using control tools View paper
- [48] The robust way to stack and bag: the local Lipschitz way View paper
- [49] Lipschitz Constants of Hybrid Zonotope Representations of Feedforward Neural Networks View paper
- [50] Efficiently Computing Local Lipschitz Constants of Neural Networks via Bound Propagation View paper
- [51] Fast and accurate optimization on the orthogonal manifold without retraction View paper
- [52] A feasible method for optimization with orthogonality constraints View paper
- [53] Online optimization over Riemannian manifolds View paper
- [54] Cross-Coupling Matrix Reconfiguration Using the LevenbergâMarquardt Algorithm on Orthogonal Groups View paper
- [55] Amortized eigendecomposition for neural networks View paper
- [56] Cheap orthogonal constraints in neural networks: A simple parametrization of the orthogonal and unitary group View paper
- [57] Quotient Geometry of Bounded or Fixed-Rank Correlation Matrices View paper
- [58] Smoothly Evolving Geodesics in the Special Orthogonal Group: Definitions, Computations and Applications View paper
- [59] Numerical Approaches for Constrained and Unconstrained, Static Optimization on the Special Euclidean Group SE (3) View paper
- [60] A randomized feasible algorithm for optimization with orthogonal constraints View paper
- [61] Robust and provably monotonic networks View paper
- [62] A closer look at accuracy vs. robustness View paper
- [63] Unlocking deterministic robustness certification on imagenet View paper