序号	首发日期	刊/会	Venue	论文名称	类型	内容	作者	单位	状态
1	20170523	C	NIPS2017	20170523-[NIPS2017]-Formal Guarantees on the Robustness of a Classifier against Adversarial Manipulation					
2	20171102	C	ICML2018	20171102-[ICML2018]-Provable Defenses against Adversarial Examples via the Convex Outer Adversarial Polytope					
3	20180129	C	ICLR2018	20180129-[ICLR2018]-Certified defenses against adversarial examples					
4	20180209	C	S&P2019	20180209-[S&P2019]-Certified robustness to adversarial examples with differential privacy					
5	20180317	C	AUAI2018	20180317-[AUAI2018]-A Dual Approach to Scalable Verification of Deep Networks					
6	20180425	C	ICML2018	20180425-[ICML2018]-Towards Fast Computation of Certified Robustness for ReLU Networks					
7	20180531	C	NIPS2019	20180531-[NIPS2019]-Scaling provable adversarial defenses					
8	20180910	C	NIPS2019	20180910-[NIPS2019]-Certified Adversarial Robustness with Additive Noise					
9	20181102	C	NIPS2018	20181102-[NIPS2018]-Efficient Neural Network Robustness Certification with General Activation Functions					
10	20181218	C	ICML2019	20181218-[ICML2019]-PROVEN-Verifying Robustness of Neural Networks with a Probabilistic Approach					
11	20190208	C	ICML2019	20190208-[ICML2019]-Certified Adversarial Robustness via Randomized Smoothing					
12	20190223	C	NIPS2019	20190223-[NIPS2019]-A Convex Relaxation Barrier to Tight Robustness Verification of Neural Networks					
13	20190315	C	ICML2019	20190315-[ICML2019]-On Certifying Non-Uniform Bounds against Adversarial Attacks					
14	20190320	C	NIPS2019	20190320-[NIPS2019]-Provable Certificates for Adversarial Examples-Fitting a Ball in the Union of Polytopes					
15	20190323	C	ICML2020	20190323-[ICML2020]-Scalable Differential Privacy with Certified Robustness in Adversarial Learning					
16	20190610	C	NIPS2019	20190610-[NIPS2019]-Robustness Verification of Tree-based Models					
17	20190612	C	NIPS2019	20190612-[NIPS2019]-Tight Certificates of Adversarial Robustness for Randomly Smoothed Classifiers					
18	20190615	C	CVPR2019	20190615-[CVPR2019]-Robustness Verification of Classification Deep Neural Networks via Linear Programming					
19	20190821	C	BigData2019	20190821-[BigData2019]-Denoising and Verification Cross-Layer Ensemble Against Black-box Adversarial Attacks					
20	20190923	C	NIPS2019	20190923-[NIPS2019]-Verified Uncertainty Calibration					
21	20191031	C	NIPS2019	20191031-[NIPS2019]-Certifiable Robustness to Graph Perturbations					
22	20191108	C	ICML2020	20191108-[ICML2020]-Certified Data Removal from Machine Learning Models					
23	20191208	C	NIPS2019	20191208-[NIPS2019]-Beyond the Single Neuron Convex Barrier for Neural Network Certification					
24	20191208	C	NIPS2019	20191208-[NIPS2019]-Certifying Geometric Robustness of Neural Networks					
25	20191219	C	CVPR2020	20191219-[CVPR2020]-Towards Verifying Robustness of Neural Networks Against A Family of Semantic Perturbations					
26	20200207	C	ICML2020	20200207-[ICML2020]-Certified Robustness to Label-Flipping Attacks via Randomized Smoothing					
27	20200208	C	ICML2020	20200208-[ICML2020]-Curse of Dimensionality on Randomized Smoothing for Certifiable Robustness					
28	20200212	C	ICLR2021	20200212-[ICLR2021]-Fast Geometric Projections for Local Robustness Certification					
29	20200221	C	ICML2020	20200221-[ICML2020]-An end-to-end approach for the verification problem learning the right distance					
30	20200221	C	NIPS2020	20200221-[NIPS2020]-Black-Box Certification with Randomized Smoothing A Functional Optimization Based Framework					
31	20200224	C	NIPS2020	20200224-[NIPS2020]-Learning Certified Individually Fair Representations					
32	20200225	C	NIPS2020	20200225-[NIPS2020]-(De)Randomized Smoothing for Certifiable Defense against Patch Attacks					
33	20200227	C	CCS2021	20200227-[CCS2021]-TSS Transformation-Specific Smoothing for Robustness Certification					
34	20200227	C	NIPS2020	20200227-[NIPS2020]-Certified Defense to Image Transformations via Randomized Smoothing					
35	20200228	C	NIPS2020	20200228-[NIPS2020]-Automatic Perturbation Analysis for Scalable Certified Robustness and Beyond					
36	20200314	C	ICLR2020	20200314-[ICLR2020]-Certified Defenses for Adversarial Patches					
37	20200319	C	ICLR2020	20200319-[ICLR2020]-Breaking Certified Defenses-Semantic Adversarial Examples with Spoofed Robustness Certificates					
38	20200507	C	NIPS2020	20200507-[NIPS2020]-Efficient Exact Verification of Binarized Neural Networks					
39	20200517	C	S&P2021	20200517-[S&P2021]-PatchGuard A Provably Robust Defense against Adversarial Patches via Small Receptive Fields and Masking					
40	20200607	C	NIPS2020	20200607-[NIPS2020]-Consistency Regularization for Certified Robustness of Smoothed Classifiers					
41	20200611	C	NIPS2020	20200611-[NIPS2020]-On the Tightness of Semidefinite Relaxations for Certifying Robustness to Adversarial Examples					
42	20200611	C	NIPS2020	20200611-[NIPS2020]-One Ring to Rule Them All Certifiably Robust Geometric Perception with Outliers					

43	20200630	C	ICML2020	20200630-[ICML2020]-Black-box Certification and Learning under Adversarial Perturbations					
44	20200713	C	ICML2020	20200713-[ICML2020]-Neural Network Control Policy Verification With Persistent Adversarial Perturbation					
45	20200716	C	NIPS2020	20200716-[NIPS2020]-Certifiably Adversarially Robust Detection of Out-of-Distribution Data					
46	20200723	C	ICML2020	20200723-[ICML2020]-Hierarchical Verification for Adversarial Robustness					
47	20200829	C	ICML2020	20200829-[ICML2020]-Efficient Robustness Certificates for Discrete Data Sparsity-Aware Randomized Smoothing for Graphs Images and More					
48	20200917	C	NIPS2020	20200917-[NIPS2020]-Certifying Confidence via Randomized Smoothing					
49	20201013	C	NIPS2020	20201013-[NIPS2020]-Higher-Order Certification For Randomized Smoothing					
50	20201022	C	NIPS2020	20201022-[NIPS2020]-Enabling certification of verification-agnostic networks via memory efficient semidefinite programming					
51	20201120	C	NIPS2020	20201120-[NIPS2020]-Certified Monotonic Neural Networks					
52	20201120	C	NIPS2020	20201120-[NIPS2020]-Certified Robustness of Graph Convolution Networks for Graph Classification under Topological Attacks					
53	20201120	C	NIPS2020	20201120-[NIPS2020]-Fast Adversarial Robustness Certification of Nearest Prototype Classifiers for Arbitrary Seminorms					
54	20201120	C	NIPS2020	20201120-[NIPS2020]-Lipschitz-Certifiable Training with a Tight Outer Bound					
55	20201120	C	NIPS2020	20201120-[NIPS2020]-Optimal Learning from Verified Training Data					
56	20201127	C	ICLR2021	20201127-[ICLR2021]-Fast and Complete Enabling Complete Neural Network Verification with Rapid and Massively Parallel Incomplete Verifiers					
57	20210114	C	ICLR2021	20210114-[ICLR2021]-Learning Safe Multi-agent Control with Decentralized Neural Barrier Certificates					
58	20210208	C	ICLR2021	20210208-[ICLR2021]-Efficient Certified Defenses Against Patch Attacks on Image Classifiers					
59	20210311	C	NIPS2021	20210311-[NIPS2021]-Beta-CROWN Efficient Bound Propagation with Per-neuron Split Constraints for Neural Network Robustness Verification					
60	20210317	C	ICML2021	20210317-[ICML2021]-Improved, Deterministic Smoothing for L_1 Certified Robustness					
61	20210401	C	CVPR2021	20210401-[CVPR2021]-Towards Evaluating and Training Verifiably Robust Neural Networks					
62	20210413	C	CVPR2021	20210413-[CVPR2021]-Simpler Certified Radius Maximization by Propagating Covariances					
63	20210503	C	ICLR2021	20210503-[ICLR2021]-Certify or Predict Boosting Certified Robustness with Compositional Architectures					
64	20210503	C	ICLR2021	20210503-[ICLR2021]-Collective Robustness Certificates Exploiting Interdependence in Graph Neural Networks					
65	20210503	C	ICLR2021	20210503-[ICLR2021]-Fooling a Complete Neural Network Verifier					
66	20210524	C	CCS2021	20210524-[CCS2021]-Learning Security Classifiers with Verified Global Robustness Properties					
67	20210602	C	NIPS2021	20210602-[NIPS2021]-Semialgebraic Representation of Monotone Deep Equilibrium Models and Applications to Certification					
68	20210701	C	ICML2021	20210701-[ICML2021]-Scalable Certified Segmentation via Randomized Smoothing					
69	20210715	C	ICML2021	20210715-[ICML2021]-CRFL-Certifiably Robust Federated Learning against Backdoor Attacks					
70	20210715	C	ICML2021	20210715-[ICML2021]-Provable Lipschitz Certification for Generative Models					
71	20210715	C	ICML2021	20210715-[ICML2021]-Towards Certifying L-infinity Robustness using Neural Networks with L-inf-dist Neurons					
72	20210715	C	ICML2021	20210715-[ICML2021]-Wasserstein Distributional Normalization For Robust Distributional Certification of Noisy Labeled Data					
73	20210804	C	NIPS2021	20210804-[NIPS2021]-Learning Barrier Certificates Towards Safe Reinforcement Learning with Zero Training-time Violations					
74	20211116	C	CCS2021	20211116-[CCS2021]-Cert-RNN Towards Certifying the Robustness of Recurrent Neural Networks					
75	20211206	C	NIPS2021	20211206-[NIPS2021]-Center Smoothing-Certified Robustness for Networks with Structured Outputs					
76	20211206	C	NIPS2021	20211206-[NIPS2021]-Certifying Robustness to Programmable Data Bias in Decision Trees					
77	20211206	C	NIPS2021	20211206-[NIPS2021]-Fast Certified Robust Training with Short Warmup					
78	20211206	C	NIPS2021	20211206-[NIPS2021]-Instance-Dependent Bounds for Zeroth-order Lipschitz Optimization with Error Certificates					
79	20211206	C	NIPS2021	20211206-[NIPS2021]-Learning Semantic Representations to Verify Hardware Designs					
80	20211206	C	NIPS2021	20211206-[NIPS2021]-Make Sure You're Unsure A Framework for Verifying Probabilistic Specifications					
81	20211206	C	NIPS2021	20211206-[NIPS2021]-ScaleCert Scalable Certified Defense against Adversarial Patches with Sparse Superficial Layers					
82	20211206	C	NIPS2021	20211206-[NIPS2021]-SmoothMix Training Confidence-calibrated Smoothed Classifiers for Certified Robustness					
83	20211206	C	NIPS2021	20211206-[NIPS2021]-Towards Better Understanding of Training Certifiably Robust Models against Adversarial Examples					
84	20211206	C	NIPS2021	20211206-[NIPS2021]-Training Certifiably Robust Neural Networks with Efficient Local Lipschitz Bounds					