

# **2025 IEEE/ACM International Workshop on Deep Learning for Testing and Testing for Deep Learning (DeepTest 2025)**

**Ottawa, Ontario, Canada  
3 May 2025**



**IEEE Catalog Number: CFP25T82-POD  
ISBN: 979-8-3315-0191-4**

**Copyright © 2025 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP25T82-POD
ISBN (Print-On-Demand):	979-8-3315-0191-4
ISBN (Online):	979-8-3315-0190-7

**Additional Copies of This Publication Are Available From:**

Curran Associates, Inc  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: (845) 758-0400  
Fax: (845) 758-2633  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2025 IEEE/ACM International Workshop on Deep Learning for Testing and Testing for Deep Learning (DeepTest) **DeepTest 2025**

## Table of Contents

Message from the Chairs .....	vii
-------------------------------	-----

### DeepTest 2025

Improving the Reliability of Failure Prediction Models through Concept Drift Monitoring .....	1
<i>Lorena Poenaru-Olaru (Software Engineering Research Group, Delft University of Technology, Netherlands), Luis Cruz (Software Engineering Research Group, Delft University of Technology, Netherlands), Jan S. Rellermeyer (Dependable and Scalable Software Systems, Leibniz University Hannover, Germany), and Arie van Deursen (Software Engineering Research Group, Delft University of Technology, Netherlands)</i>	
DANDI: Diffusion as Normative Distribution for Deep Neural Network Input .....	9
<i>Somin Kim (KAIST)</i>	
Lachesis: Predicting LLM Inference Accuracy using Structural Properties of Reasoning Paths.....	17
<i>Naryeong Kim (Korea Advanced Institute of Science and Technology)</i>	
Robust Testing for Deep Learning using Human Label Noise .....	21
<i>Gordon Lim (University of Michigan, USA), Stefan Larson (Vanderbilt University, USA), and Kevin Leach (Vanderbilt University, USA)</i>	
DILLEMA: Diffusion and Large Language Models for Multi-Modal Augmentation .....	29
<i>Luciano Baresi (Politecnico di Milano, Italy), Davide Yi Xian Hu (Politecnico di Milano, Italy), Muhammad Irfan Mas'udi (Politecnico di Milano, Italy), and Giovanni Quattrocchi (Politecnico di Milano, Italy)</i>	
Reinforcement Learning from Automatic Feedback for High-Quality Unit Test Generation .....	37
<i>Benjamin Steenhoek (Microsoft, USA), Michele Tufano (Google, USA), Neel Sundareshan (Microsoft, USA), and Alexey Svyatkovskiy (Google DeepMind, USA)</i>	
On the Effectiveness of LLMs for Manual Test Verifications .....	45
<i>Myron Peixoto (Federal University of Alagoas, Brazil), Davy Baía (Federal University of Alagoas), Nathalia Nascimento (Penn State University), Paulo Alencar (University of Waterloo), Balduino Fonseca (Federal University of Alagoas), and Márcio Ribeiro (Federal University of Alagoas)</i>	

OpenCat: Improving Interoperability of ADS Testing ..... 53  
    *Qurban Ali (University of Milano-Bicocca, Milan, Italy), Andrea Stocco*  
    *(Technical University of Munich, fortiss GmbH Munich, Germany),*  
    *Leornardo Mariani (University of Milano-Bicocca, Milan, Italy), and*  
    *Oliviero Riganelli (University of Milano-Bicocca, Milan, Italy)*

**Author Index** ..... 61