

THÈSE DE DOCTORAT DE

NANTES UNIVERSITÉ

ÉCOLE DOCTORALE N° 641
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Informatique*

Par

Yejing XIE

Handwritten Mathematical Expression Recognition with Graph Neural Networks and Tree-based Language Models

En vue d'une soutenance à Nantes, le 12 Décembre 2025

Unité de recherche : Laboratoire des Sciences du Numérique de Nantes (LS2N)

Rapporteurs avant soutenance :

M. Josep LLADÓS Professeur, Universitat Autònoma de Barcelona
Mme. Véronique EGLIN Professeure, INSA de Lyon

Composition du Jury :

Examineurs :	M. Richard ZANIBBI	Professeur, Rochester Institute of Technology
	M. Eric ANQUETIL	Professeur, INSA de Rennes
	M. Jean-Baptiste FASQUEL	Professeur, Université d'Angers
Dir. de thèse :	M. Harold MOUCHÈRE	Professeur, Nantes Université

Titre : Reconnaissance d'expressions mathématiques manuscrites à l'aide de réseaux neuronaux graphiques et de modèles linguistiques arborescents

Mot clés : Reconnaissance d'expressions mathématiques manuscrites, réseaux neuronaux graphiques, étiquetage de graphes, transformateur à structure arborescente, Modèle linguistique arborescent

Résumé : La reconnaissance des expressions mathématiques manuscrites (HMER) vise à traduire les formules écrites à la main en représentations lisibles par machine, reliant ainsi la perception visuelle au raisonnement symbolique. Malgré les progrès de l'apprentissage profond, la HMER reste un défi en raison de l'ambiguïté de l'écriture manuscrite, de la complexité spatiale bidimensionnelle et de la syntaxe mathématique complexe. Cette thèse aborde la HMER sous un angle structural en modélisant les expressions mathématiques sous forme de graphes et d'arbres en 2D. Nous proposons un réseau d'attention sur

graphes pondérés (EGAT) pour la classification conjointe des symboles et des relations, ainsi qu'un décodeur transformeur structuré en arbre (TST) pour la prédiction hiérarchique des structures. Pour l'évaluation, la compétition CROHME 2023 a été organisée avec de nouvelles données manuscrites et synthétiques à grande échelle. Les approches proposées obtiennent des performances compétitives et apportent de nouvelles perspectives et ressources pour une reconnaissance des expressions mathématiques plus structurée et sensible au contexte.

Title: Handwritten Mathematical Expression Recognition with Graph Neural Networks and Tree-based Language Models

Keywords: Handwritten Mathematical Expression Recognition, Graph Neural Networks, Graph Labeling, Tree-structured Transformer, Tree Language Model

Abstract: Handwritten Mathematical Expression Recognition (HMER) aims to translate handwritten formulas into machine-readable representations, bridging visual perception and symbolic reasoning. Despite progress in deep learning, HMER remains challenging due to handwriting ambiguity, 2D spatial complexity, and intricate mathematical syntax. This thesis addresses HMER from a structural perspective by modeling mathematical expressions as graphs and trees in 2D structure. We propose an Edge-weighted Graph Attention Network (EGAT) for joint symbol and

relation classification, and a Tree-Structured Transformer (TST) Decoder for hierarchical structure prediction. To facilitate evaluation, the CROHME 2023 competition was organized with newly collected handwritten and large-scale synthetic data. Contextual extensions of CROHME were further introduced to assess the role of textual and visual context using vision-language models. The proposed approaches achieve competitive performance and contribute new insights and resources toward more structured and context-aware mathematical expression recognition.