# CircuitFusion: Multimodal Circuit Representation Learning for Agile Chip Design

香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

About Me

**Wenji Fang** Shang Liu Jing Wang Zhiyao Xie

Hong Kong University of Science and Technology

Paper Code

ICLR International Conference On Learning Representations

## Highlights

**Why?** Most AI for EDA solutions are task-specific and overlook the multimodal nature of circuits.

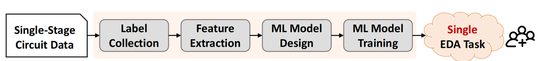
**What?** Learn general circuit embeddings from multimodalities, supporting various EDA tasks and outperforming task-specific methods.

**How?** Identify unique circuit properties and propose tailored strategies to build a multimodal, implementation-aware RTL circuit encoder.

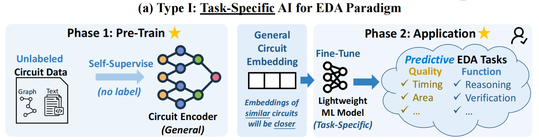
## Introduction

### ❑ Paradigm Shift of AI for EDA

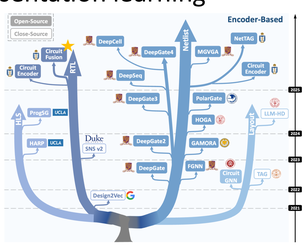
- Widely adopted: task-specific supervised learning
- New trend: general self-supervised representation learning




(a) Type I: Task-Specific AI for EDA Paradigm



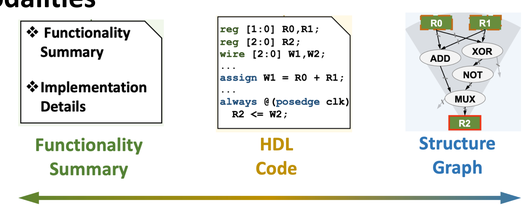

(b) Type II: General Encoder-Based Circuit Foundation Model Paradigm




### ❑ Limitation: only encode circuit *graph structure*

## Circuit Multimodality

### ❑ 3 RTL modalities

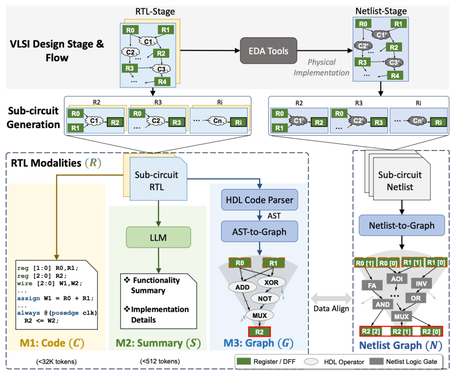



## Method

### ❑ Key Idea: 4 unique circuit properties (P1-P4) → 4 tailored strategies (S1-S4)

### ❑ Model Architecture:

- CircuitFusion
  - 3 unimodal encoders
  - multimodal fusion encoder
- Auxiliary netlist encoder




### ❑ Step 1. *Preprocessing*

[P1] Parallel execution

[S1] Sub-circuit generation

- Split via register cones

### ❑ Step 2. *Pre-training*

[P2] Functional equivalence

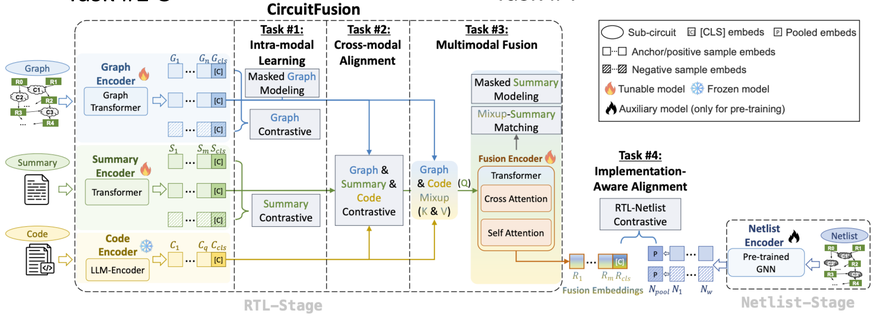
[S2] Semantic-structure pre-training

- Task #1-3

[P3] Multiple design stages

[S3] Implementation-aware alignment

- Task #4

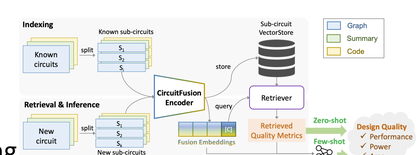



### ❑ Step 3. *Application*

[P4] Circuit Reusability

[S4] Retrieval-augmented infernce

- Retrieves most similar circuits
- Enable zero-shot & improve fine-tuning




## Experimental Results

### ❑ Comparison w. Baselines

- Various tasks: slack, WNS, TNS, power, area prediction
- Outperforming task-specific / text / software solutions

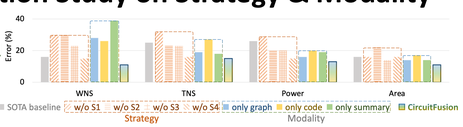
| Type | Method | Slack R | Slack MAPE | WNS R | WNS MAPE | TNS R | TNS MAPE | Power R | Power MAPE | Area R | Area MAPE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Hardware Solution | RTL-Timer | 0.85 | 17% | 0.9 | 16% | 0.96 | 25% | N/A | | N/A | |
| | MasterRTL | N/A | | 0.89 | 18% | 0.94 | 28% | 0.89 | 26% | 0.98 | 16% |
| | SNS v2 | N/A | | 0.82 | 22% | N/A | | 0.76 | 28% | 0.93 | 25% |
| Text Encoder | NV-Embed-v1 | N/A | | 0.49 | 17% | 0.97 | 55% | 0.85 | 44% | 0.86 | 24% |
| Software Code Encoder | UnixCoder | N/A | | 0.46 | 21% | 0.95 | 44% | 0.83 | 29% | 0.85 | 26% |
| | CodeT5+ Encoder | N/A | | 0.55 | 21% | 0.63 | 43% | 0.49 | 46% | 0.45 | 39% |
| | CodeSage | N/A | | 0.23 | 25% | 0.86 | 45% | 0.8 | 38% | 0.77 | 41% |
| **Ours** | **CircuitFusion** | **0.87** | **12%** | **0.91** | **11%** | **0.99** | **15%** | **0.99** | **13%** | **0.99** | **11%** |

### ❑ Zero-shot Prediction via Retrieval

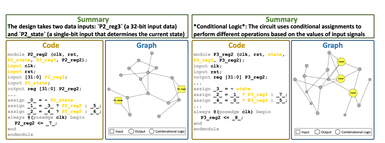
Table 3: MAPE(%) results of the zero-shot top-k similar circuit retrieval.

| Method | Slack top-1 | Slack top-3 | Slack top-5 | Slack top-10 | Sub-circuit Power top-1 | Sub-circuit Power top-3 | Sub-circuit Power top-5 | Sub-circuit Power top-10 | Sub-circuit Area top-1 | Sub-circuit Area top-3 | Sub-circuit Area top-5 | Sub-circuit Area top-10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LLM Encoder | 51 | 35 | 33 | 34 | 92 | 90 | 90 | 90 | 90 | 88 | 88 | 88 |
| UnixCoder | 56 | 36 | 36 | 36 | 90 | 89 | 90 | 91 | 89 | 88 | 89 | 89 |
| CodeT5+ Embedding | 57 | 35 | 35 | 36 | 88 | 87 | 89 | 90 | 87 | 86 | 87 | 88 |
| CodeSage | 50 | 36 | 36 | 36 | 89 | 87 | 88 | 91 | 88 | 85 | 86 | 87 |
| **Ours** | **21** | **22** | **23** | **26** | **36** | **40** | **42** | **53** | **35** | **40** | **42** | **51** |

### ❑ Ablation Study on Strategy & Modality

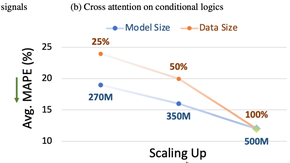



### ❑ Visualization of Multimodal Cross Attention




(a) Cross attention on input signals (b) Cross attention on conditional logics

### ❑ Scaling w. Model / Data Size




## Conclusion & Future Work

### ❑ Conclusion: first general multimodal RTL encoder

### ❑ Future Work

- Multimodal netlist encoder [*DAC'25*]
- Align circuit encoders with generative LLM decoders