ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

XXXV Cycle

A Graph Neural Network based High-Level Synthesis Performance Prediction

M. Usman Jamal
Supervisor: Prof. Luciano Lavagno

Research context and motivation

- Traditional Electronic Design Automation (EDA) tools usually take hours to days to accurately evaluate the circuit quality with considerable manual effort. While these hardware designs are efficient, they also require extensive design-space exploration (DSE).

- High-level Synthesis (HLS) tools can significantly reduce the design effort, enabling a faster DSE but HLS-estimated circuit quality can be highly inaccurate with respect to the actual quality of results (QoR). These inaccurate estimates force the designers to turn to the time-consuming downstream implementation tasks to get actual QoR.

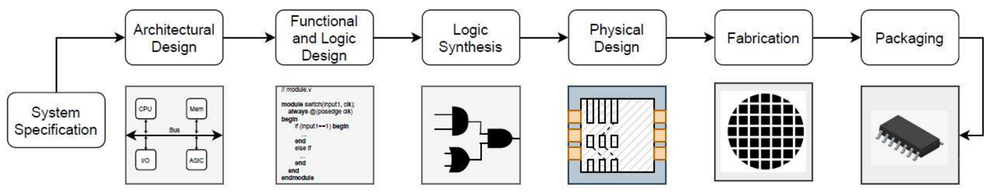

Fig. 1 – Chip Design Flow [1]

[1] Daniela et. al, '' A Survey of Graph Neural Networks for Electronic Design Automation '', MLCAD 2022.

Addressed research questions/problems

- Problem: **Fast and Accurate Performance Prediction**
  - Traditional Tools - Accurate but slow.
  - HLS - Relatively fast but still requires synthesis and implementation to accurately evaluate QoR.

- Current State-of-the-art Research

	ML Model		Task Details		C-Synthesis Required	Ground Truth (Ref. Labels)	TARGET	
	Graphical	Non-Graphical	Resource Usage	Delay			FPGA	ASIC
[1]		✓	✓	✓	Yes	Post-Route	✓	
[2]	✓			✓	Yes	Post-Route	✓	
[3]		✓	✓	✓	Yes	Post-Route	✓	
[4]	✓		✓	✓	Scheduled DFG	Post-Route	✓	
[5]	✓	✓		✓	Yes	Logic Synthesis		✓
This Work	✓		✓	✓	No	Post-Route	✓	

Table 1 - ML-based HLS Prediction Research

- Current approaches for FPGAs are done for old tools like Vivado HLS and most of the works are not open-source. Plus, they also require to go through the full HLS process which can be time-consuming.

[1] S. Dai et. al, " Fast and accurate estimation of quality of results in high-level synthesis with machine learning ", FCCM 2018.
[2] E. Ustun et al., " Accurate Operation Delay Prediction for FPGA HLS Using Graph Neural Networks ", ICCAD 2020.
[3] H. Makrani et. al, " Pyramid: Machine learning framework to estimate the optimal timing and resource usage of a high-level synthesis design ", FPL 2019.
[4] N. Wu et. al, " Ironman: GNN-assisted design space exploration in high-level synthesis via reinforcement learning ", ACM GLSVLSI 2021.
[5] S. De et. al, " Delay Prediction for ASIC HLS: Comparing Graph-based and Non-Graph-based Learning Models ", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022.

Novel contributions

- **An open-source C/C++-to-FPGA Dataset**
  - Real-world designs from diverse benchmarks
  - Each design run through complete C/C++-to-bitstream
  - Implemented using various pragmas/directives
  - With different Clock Periods
  - Using state-of-the-art tool (Vitis HLS)

- **Our Proposed Solution**

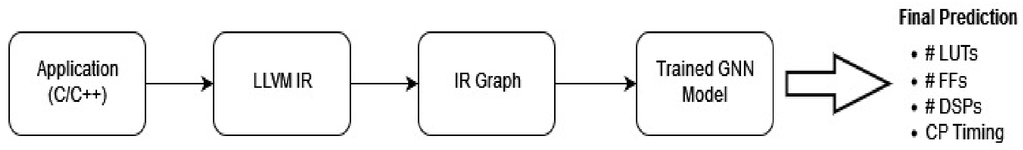

Fig. 2 – Prediction/Inference Flow

Adopted methodologies

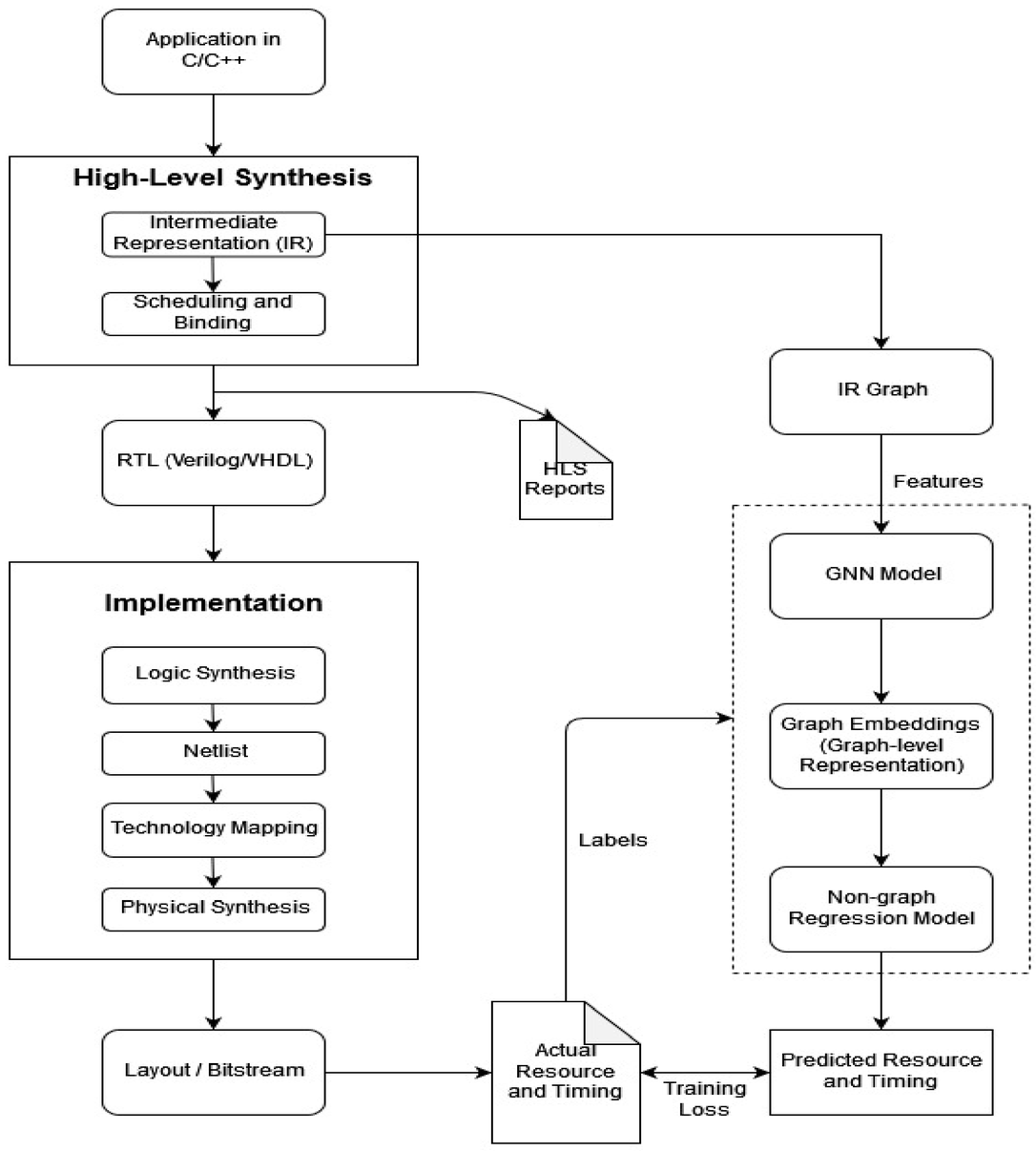

Fig. 3 – Training working flow

Future work

- Investigation of the state-of-the-art GNN architectures to exploit their representation power to predict the actual performance of the circuit expected after the implementation phase.

- Development of a GNN based model to do fast and accurate QoR prediction.

List of attended classes

- 01UJBRV - Adversarial training of neural networks (1/7/20, 15h)
- 01QTEIU - Data mining concepts and algorithms (1/2/2021, 20h)
- 01TVUQW - Embedded Electronic Systems for AI/ML (14/9/2021, 30h)
- 01QSAIU - Heuristics and metaheuristics for problem solving: new trends and software tools (10/7/20, 20h)
- 01MNFIU - Parallel and distributed computing (22/7/20, 25h)
- 01QSCIU - Reconfigurable computing (20/7/20, 20h)
- 01DUCRV - Principles of digital image processing and technologies (22/7/22, 27h)
- 01UNRRV - Entrepreneurship and start-up creation (3/7/20, 40h)
- 01UNWRV - Intercultural & interpersonal management (3/6/21, 8h)
- 01QORRV - Writing Scientific Papers in English (25/6/20, 15h)
- 01UNURV - Project management II (24/6/21, 8h)

External Research Activity

- Inria-DFKI European Summer School on A.I. (IDESSAI 2022)

Submitted and published works

- (Submitted) Brignone G., Jamal M. U., Lazarescu M. T., Lavagno L., "Array-specific dataflow caches for high-level synthesis of memory-intensive algorithms on FPGAs", IEEE Access.

POLITECNICO DI TORINO

PhD program in
Electrical, Electronics and Communications Engineering