ICCR 2025

Efficient Grasp Detection via Knowledge Distillation: A Lightweight Generative Grasping Convolutional Neural Network Framework

Presenters: Jie-Yu Chao, Po-Lien Wu, Jr-Yu Chang

Authors: Sheng-Kai Chen, Jie-Yu Chao, Po-Lien Wu, Jr-Yu Chang and Po-Chiang Lin

Date: December 5th, 2025

CONTENTS

MOTIVATION

Presenting the motivation for addressing these issues.

Problems

- High computational cost of deep grasp detection models limits deployment on embedded systems or edge devices.
- Original GGCNN, though efficient, still struggles with real-time inference under limited resources.

Goal

Achieve lightweight, real-time GGCNN models that maintain comparable IoU accuracy with much smaller size and faster speed.

Compression Comparison

- **Pruning:**
  Remove unnecessary weights
  → smaller and faster, but may lose accuracy
- **Quantization:**
  Use lower-precision numbers
  → saves memory, but hardware-dependent
- **Knowledge Distillation:**
  Train a small student to learn from a large teacher
  → keeps good accuracy with less size

CONTRIBUTIONS

Showing why this research is important to the field.

KD-based GGCNN Compression Framework

Tailored KD framework preserves dense grasp prediction while reducing model complexity.

Enhanced Deployment Efficiency

Faster inference, reduced memory usage, and higher throughput with minimal accuracy loss.

Lightweight Student Architectures

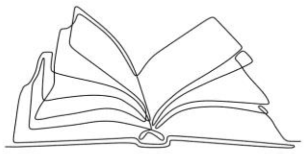
Two optimized student models exploring trade-offs in speed, size, and accuracy.

Real-time Edge Application

Demonstrated strong performance on resource-limited robotic platforms.

RELATED WORK

Reviewing existing methods and research progress.



Point-wise Knowledge Distillation

- Pixel-level alignment of teacher and student outputs
- Soft-target KD
- Enables lightweight GGCNN grasp prediction

Reference: D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv preprint arXiv:1503.02531*, 2015

Structural Knowledge Distillation

- Preserves spatial & relational consistency
- Effective for segmentation and dense tasks
- Future extension for grasp detection

Reference: L. Peng, R. Cai, J. Xiang, J. Zhu, W. Liu, W. Gao, and Y. Liu, "LiteGrasp: A Light Robotic Grasp Detection via Semi-Supervised Knowledge Distillation," *IEEE Robotics and Automation Letters*, vol. 9, pp. 7995–8002, 2024

Knowledge Distillation with Multiple Teachers

- Aggregates diverse knowledge for robustness
- High training cost in prior studies
- This work adopts single-teacher GGCNN for efficiency

Reference: Y. Liu, K. Li, P. Sun, Y. Zhang, and C. Li, "Structured knowledge distillation for semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019, pp. 2604–2613.

Real-Time Robotic Deployment and Edge Devices

- Edge devices require compact, low-latency models
- KD balances accuracy & efficiency
- Two student models achieve faster, smaller, real-time performance

Reference: H. Li, K. Ota, and M. Dong, "Learning IoT in Edge: Deep Learning for the Internet of Things with Edge Computing," *IEEE Network*, vol. 32, no. 1, pp. 96–101, Jan. 2018.

KNOWLEDGE DISTILLATION FOR GGCNN

The structure of Knowledge Distillation.

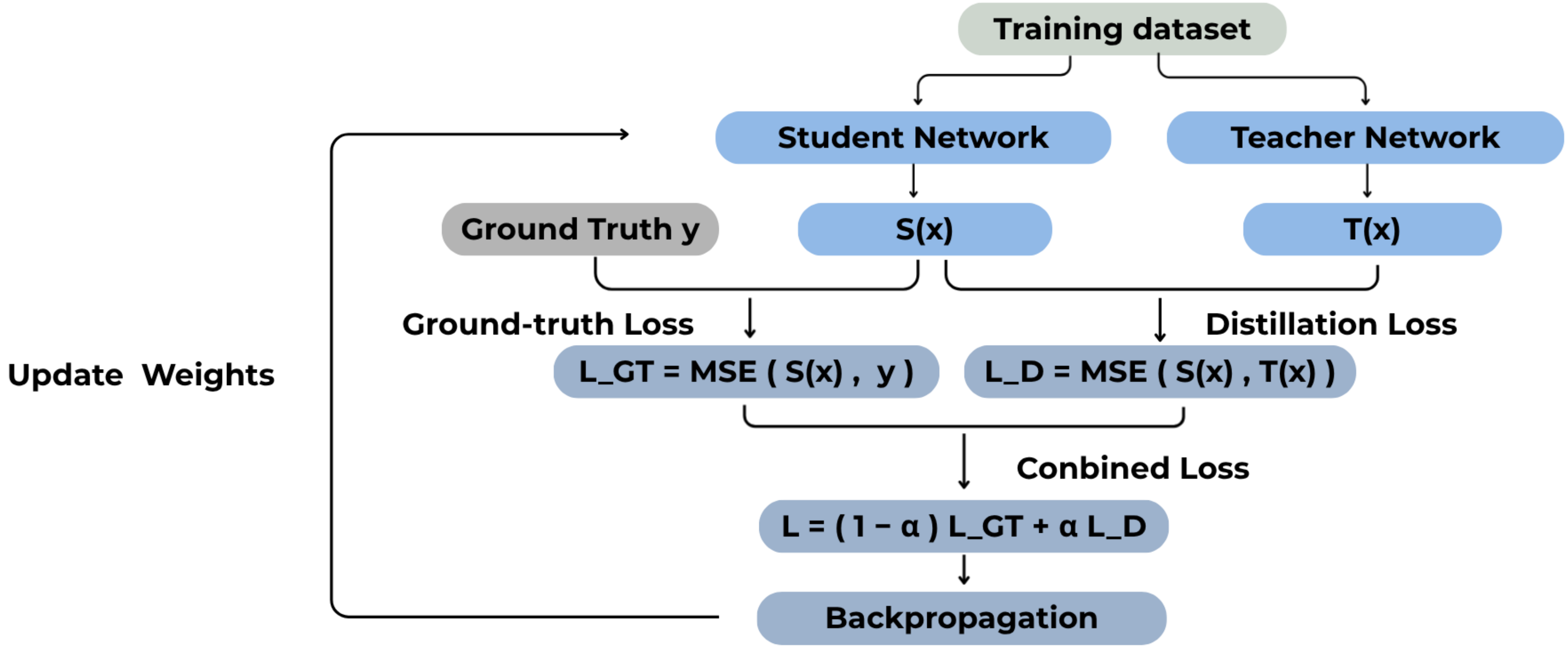


Fig 1. Knowledge Distillation Framework

KNOWLEDGE DISTILLATION FOR GGCNN

The structure of Knowledge Distillation Framework.

Performance Gap Reduction

Structured KD effectively narrows the performance gap between teacher and student models.

Stable Convergence

Progressive supervision enhances stability and ensures reliable convergence.

Soft Targets

The pre-trained, frozen teacher model provides soft targets for distillation.

Architectural Flexibility

No layer sharing—output-only distillation enables flexible and efficient student model design.

Hybrid Loss

The student model is trained end-to-end with a weighted MSE combining ground truth and teacher outputs.

LOSS COMPUTATION AND OPTIMIZATION

Describing the formulation and components of the loss function.

Loss Formulation

- **Hybrid loss:**
  combination of ground truth and distillation losses
- **Equation:**
  $L = (1 - \alpha) \cdot L_{GT} + \alpha \cdot L_{KD}$

Loss Components

- **Ground Truth Loss (L_{GT}):**
  MSE between student outputs and labeled data
  1. grasp quality p
  2. angle components c = cos(θ), s = sin(θ)
  3. grasp width w
- **Distillation Loss (L_{KD}):**
  MSE between student and teacher outputs on the same attributes

LOSS COMPUTATION AND OPTIMIZATION

Highlighting the relationship between loss computation and training benefits.

Mean Squared Error (MSE)

a_i: model prediction

b_i: target value

N: number of prediction elements

Training Benefits

- Dual-objective supervision enhances feature learning
- Faster convergence and better generalization
- Effective under resource-constrained environments

$$MSE(a,b) = \frac{1}{N}\sum_{i=1}^{n}(a_i - b_i)^2$$

Fig 2. Mean Squared Error Formula

STUDENT MODELS

The structure of each models.

Student GGCNN v1

- **Encoder** :

  Conv (16, k = 9, s = 3) → Conv (8, 5, 2) → Conv (4, 3, 2)

- **Decoder** :

  DeConv (4, 3, 2) → DeConv (8, 5, 2) → DeConv (16, 9, 3)

- **Output Heads** : p, c, s, w = Conv2D(h_6, 1, 2)

Student GGCNN v2

- **Encoder** :

  Conv(24, k = 9, s = 3) → Conv(12, 5, 2) → Conv(6, 3, 2)

- **Decoder** :

  DeConv(6, 3, 2) → DeConv(12, 5, 2) → DeConv(24, 9, 3)

- **Output Heads** : p, c, s, w = Conv2D(h_6, 1, 2)

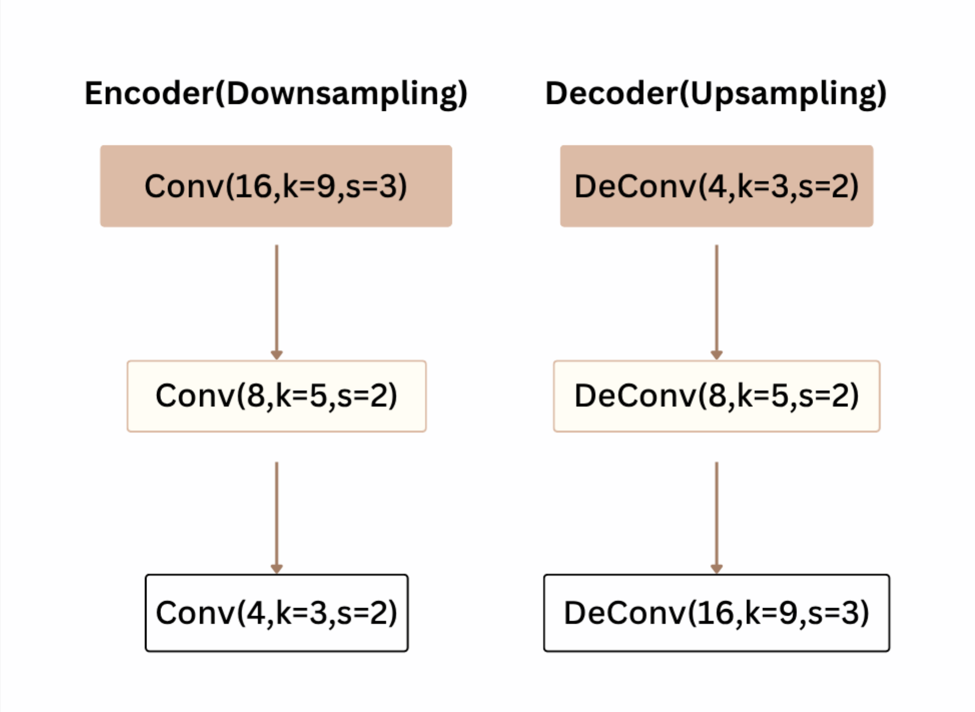


Fig 3. Student Model v1 Structure

GRASP POST-PROCESSING

Presenting how post-processing steps prepare grasp outputs for evaluation.

Evaluation

Measured inference speed (FPS) and grasp accuracy.

Key Steps

- Normalize quality (Sigmoid)
- Ensure positive width (ReLU)
- Compute final angle

Post-processing

Converts raw model outputs into final predictions.

- Quality (p)
- Width (w)
- Angle components (cosine c and sine s)

Optional

Filtering to remove noise and smooth results.

IOU-BASED GRASP SUCCESS EVALUATION

Defining success and key metrics for IoU-based grasp evaluation.

Success Definition

A grasp is "successful" if its IoU (Intersection over Union) with a ground truth box is 0.25%

Key Performance Metrics

- Grasp success rate (%)
- Average inference time (ms)
- Inference speed (FPS)
- Model parameter count

Goal

To systematically assess the real-time efficiency and prediction accuracy of the studaent models.

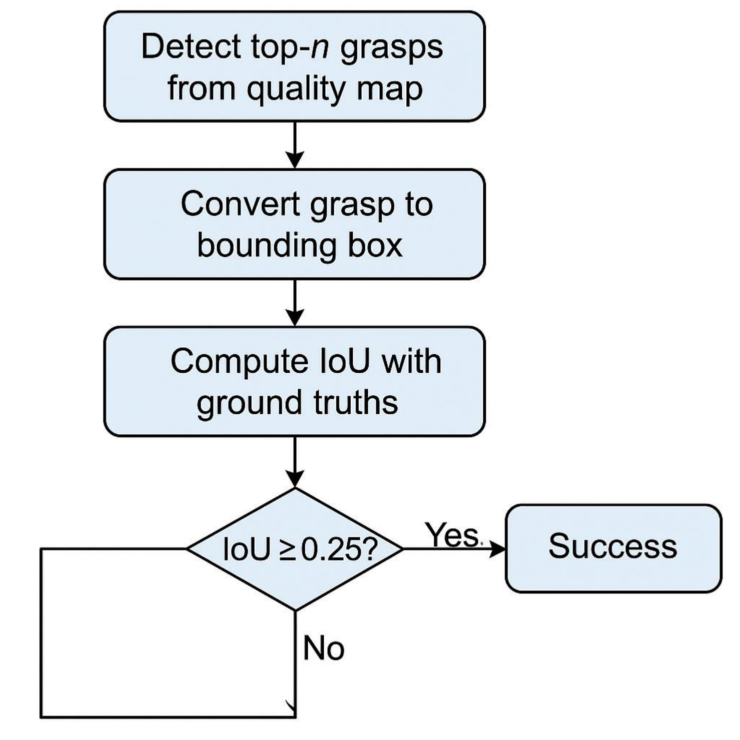


Fig4. Evaluation Flow

RESULTS

Comparison between various models.

Model Size

- **Teacher (GGCNN):** 0.24 MB
- **Student v1 (Small):** 0.06 MB → −75%
- **Student v2 (Medium):** 0.14 MB → −42%

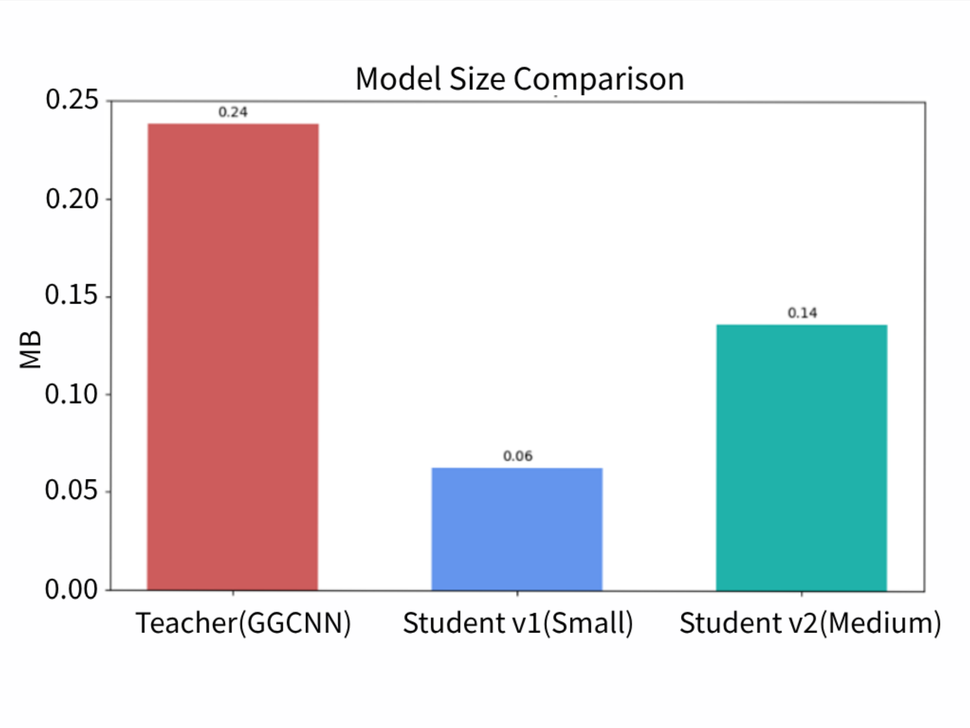


Fig5. Model Size Comparison

RESULTS

Comparison between various models.

Inference Speed

- **Teacher:** 21.74 ms
- **Student v1:** 6.99 ms (≈ 3× faster)
- **Student v2:** 9.30 ms (≈ 2.3× faster)

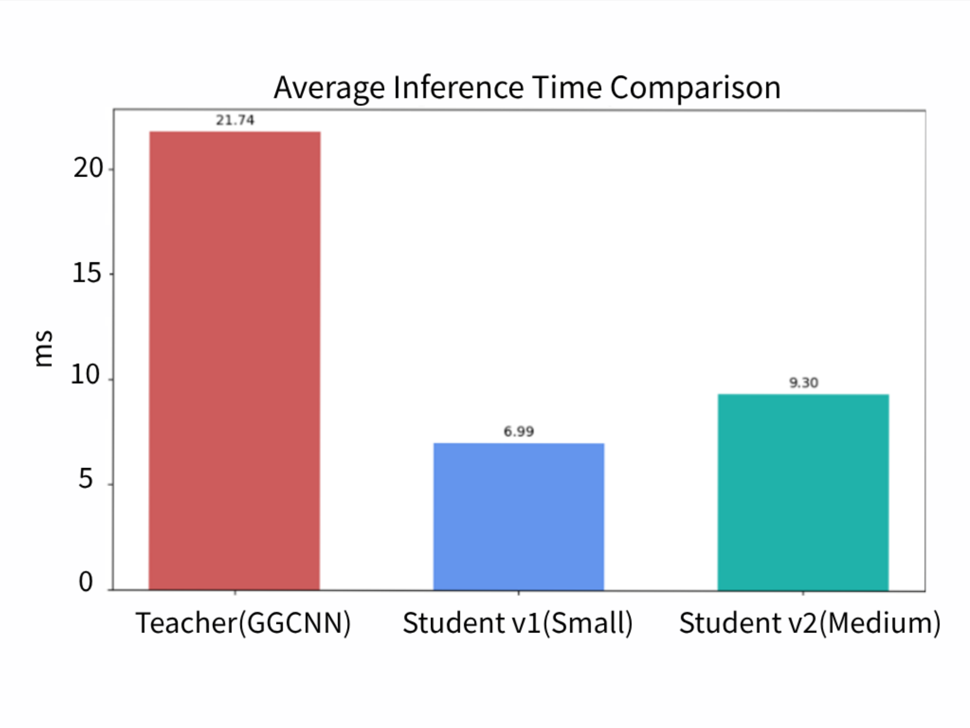


Fig6. Average Inference Time Comparison

RESULTS

Comparison between various models.

Accuracy (IoU)

- **Teacher:** 83.1 %
- **Student v1:** 79.8 % (–3.3 % drop)
- **Student v2:** 83.1 % = same as teacher

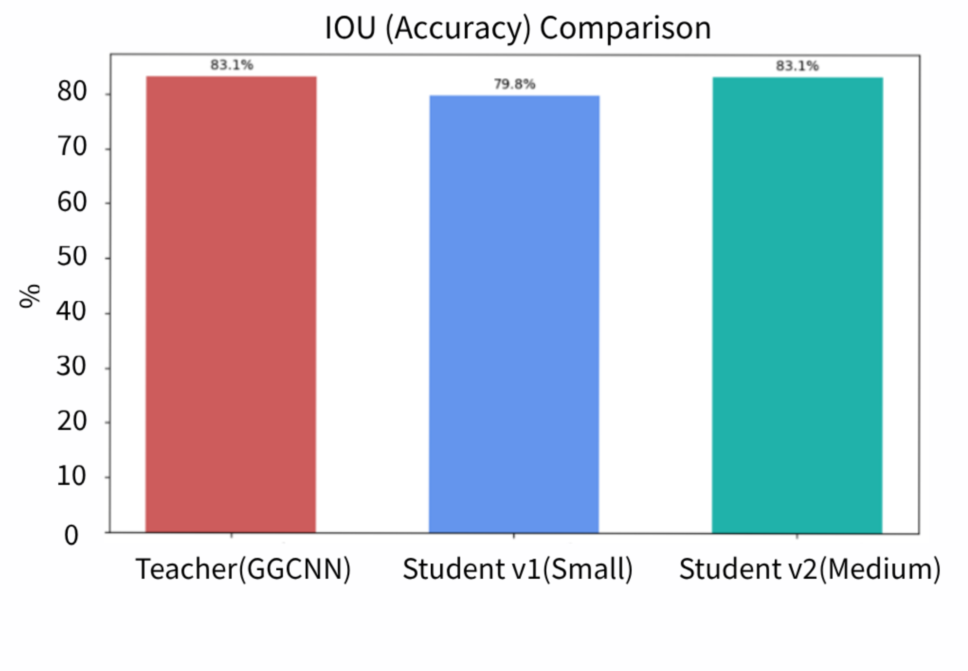


Fig7. Model Size Comparison

RESULTS

Comparison between various models.

Throughput (FPS)

- **Teacher:** 46.0 FPS
- **Student v1:** 143.1 FPS (↑ 3×)
- **Student v2:** 107.5 FPS (↑ 2.3×)

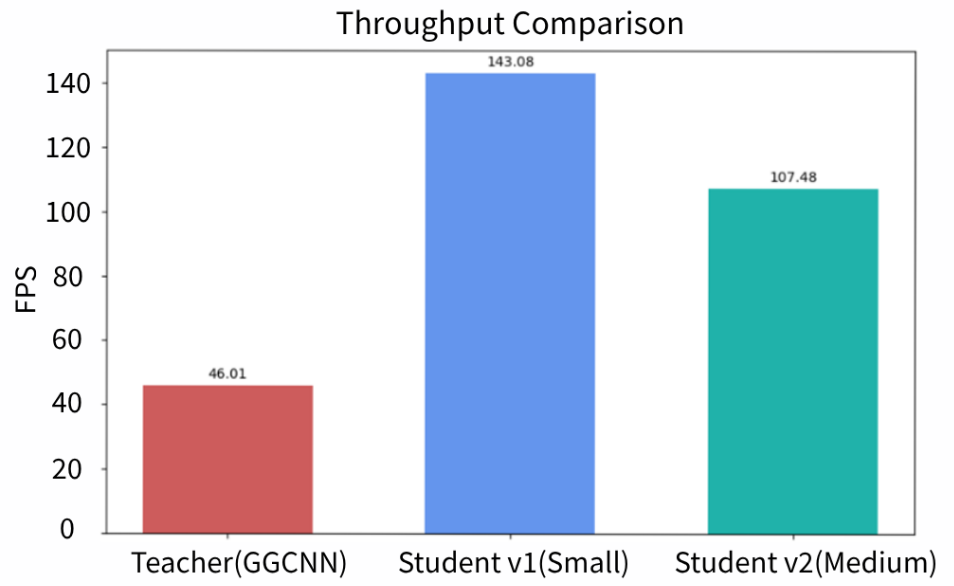


Fig8. Average Inference Time Comparison

CONCLUSION

Summary of Results and Future Extensions

Summary

- **Proposal:**
  A Knowledge Distillation (KD) framework to compress the GGCNN model.
- **Result:**
  Created two student models balancing speed, size, and accuracy.
  Student v1 (Speed): Prioritizes speed and small size.
  Student v2 (Accuracy): Maintains teacher's accuracy but is faster.
- **Key Value:**
  An easy way to get real-time grasping on resource-constrained hardware.

Future work

- **Explore**:
  Advanced KD (structural, multi-teacher)
  Hybrid methods (KD + pruning/quantization)
- **Extend:**
  Apply the framework to richer inputs like RGB-D or multi-view data.
- **Challenge:**
  Test models in more complex scenarios, like cluttered environments or dynamic tasks.

Thank You

Jie-Yu Chao

s1120302@mail.yzu.edu.tw

Po-Lien Wu

s1120447@mail.yzu.edu.tw

Jr-Yu Chang

s1120437@mail.yzu.edu.tw

Sheng-Kai Chen

s1134807@mail.yzu.edu.tw

Po-Chiang Lin

pclin@saturn.yzu.edu.tw