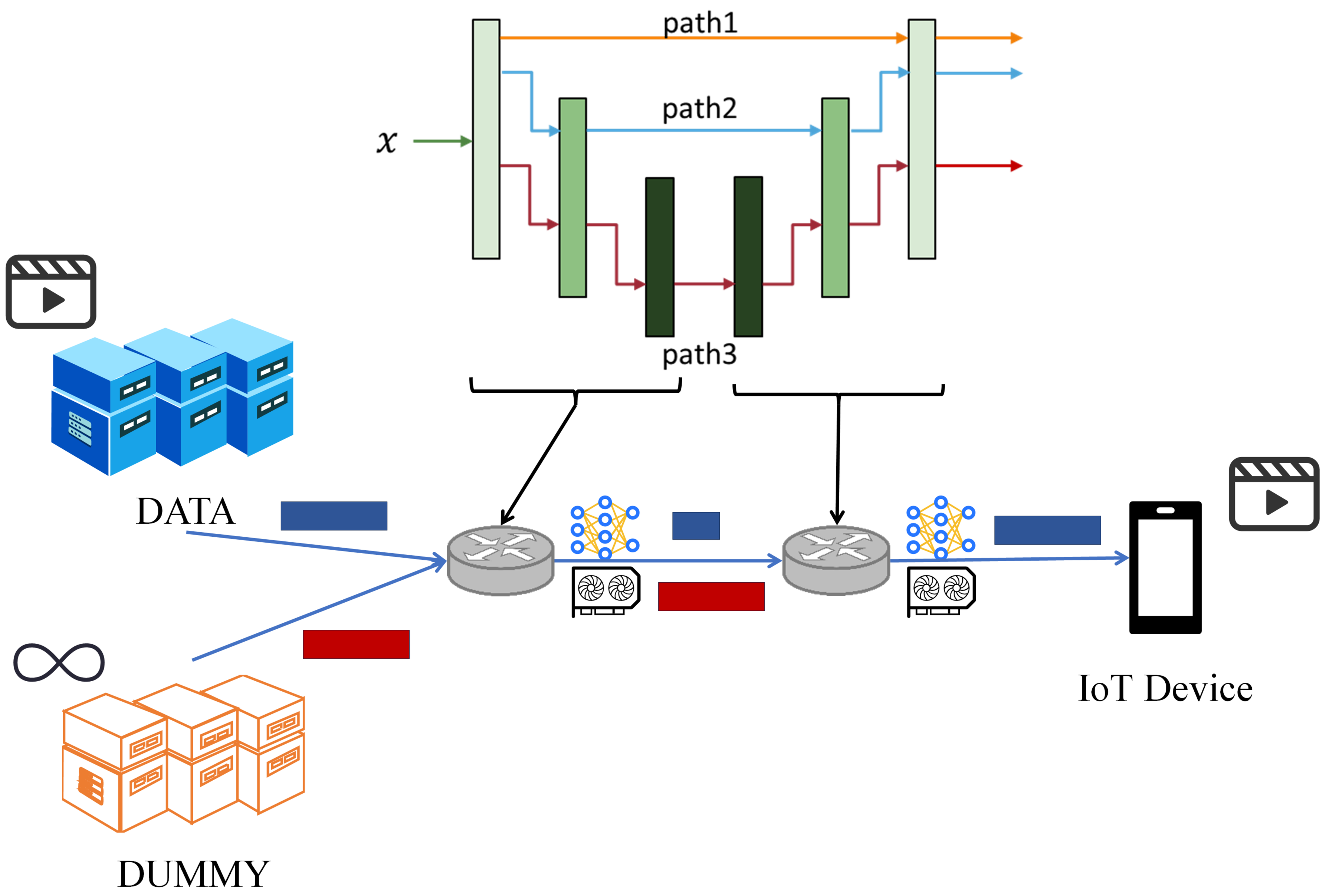
RAINLAB
Reinforcing AI & Network Lab

# Dynamic Neural Network in Computer Networks: Enhancing Network Efficiency

Junhao Cai, Changhee Joo
Computer Science and Engineering, Korea University, Seoul 02841, South Korea

## Introduction

- **In-network compression:** Video traffic requires significant bandwidth or end-device computation; lacking these leads to low QoS. In-network compression serves as an alternative, balancing network and computational resources.
- **Feedback-Based Multimedia Transmission:** Previous methods depend on feedback about resource availability, often leading to inaccurate decisions due to stale or inaccurate data.
- **Dynamic Autoencoder:** We propose a dynamic neural network-based compression method enabling flexible in-network computation without the need for feedback.

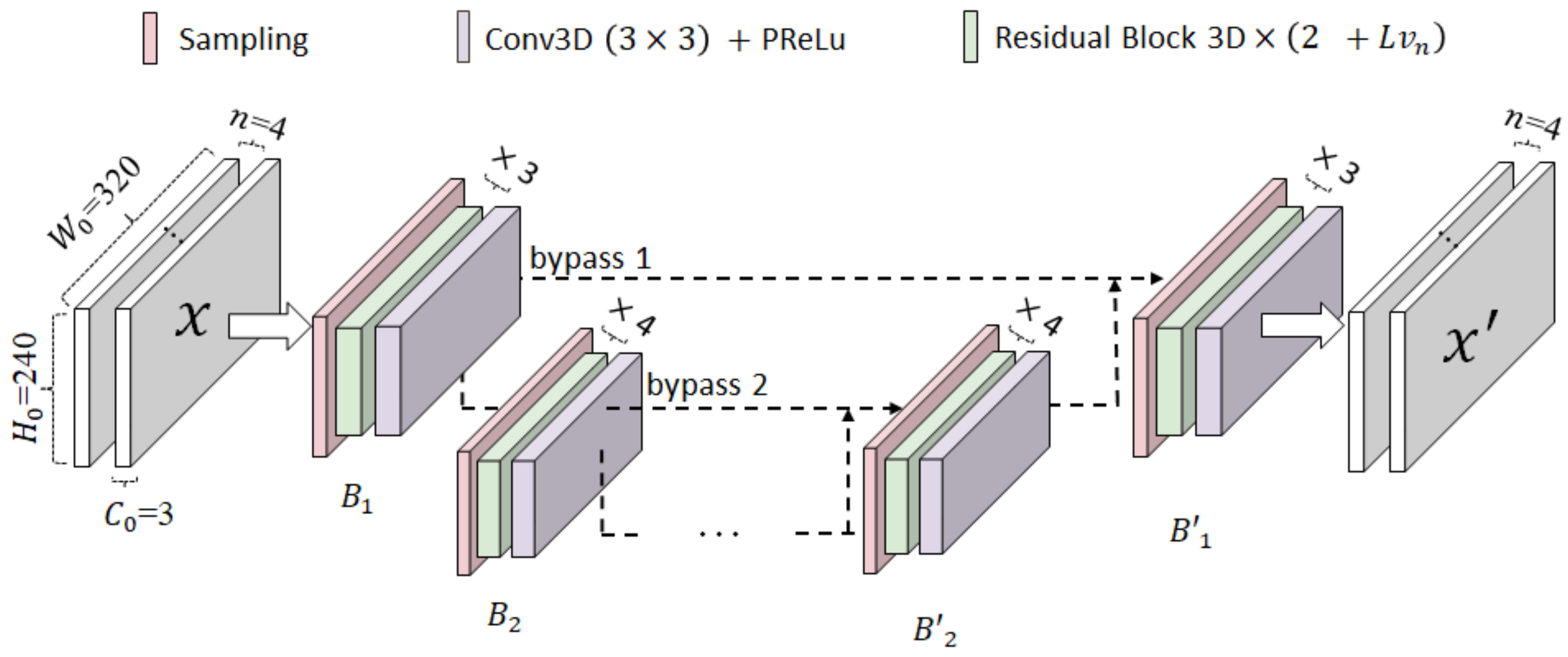



## Objectives

- **Model Structure:** Develop a dynamic autoencoder model utilizing dynamic neural network technology.
- **Flexible Compression Ratio:** Adjust the compression ratio via multi-level compression.
- **Network Resource Usage:** Enable flexible utilization of network resources to adapt to dynamic network environments.
- **Better Service Quality:** Address the decline in service quality due to increasing multimedia data traffic.

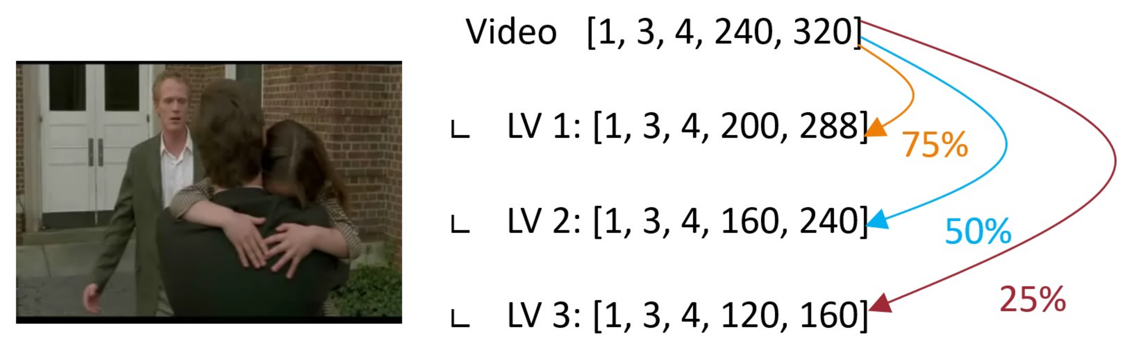
## Methods

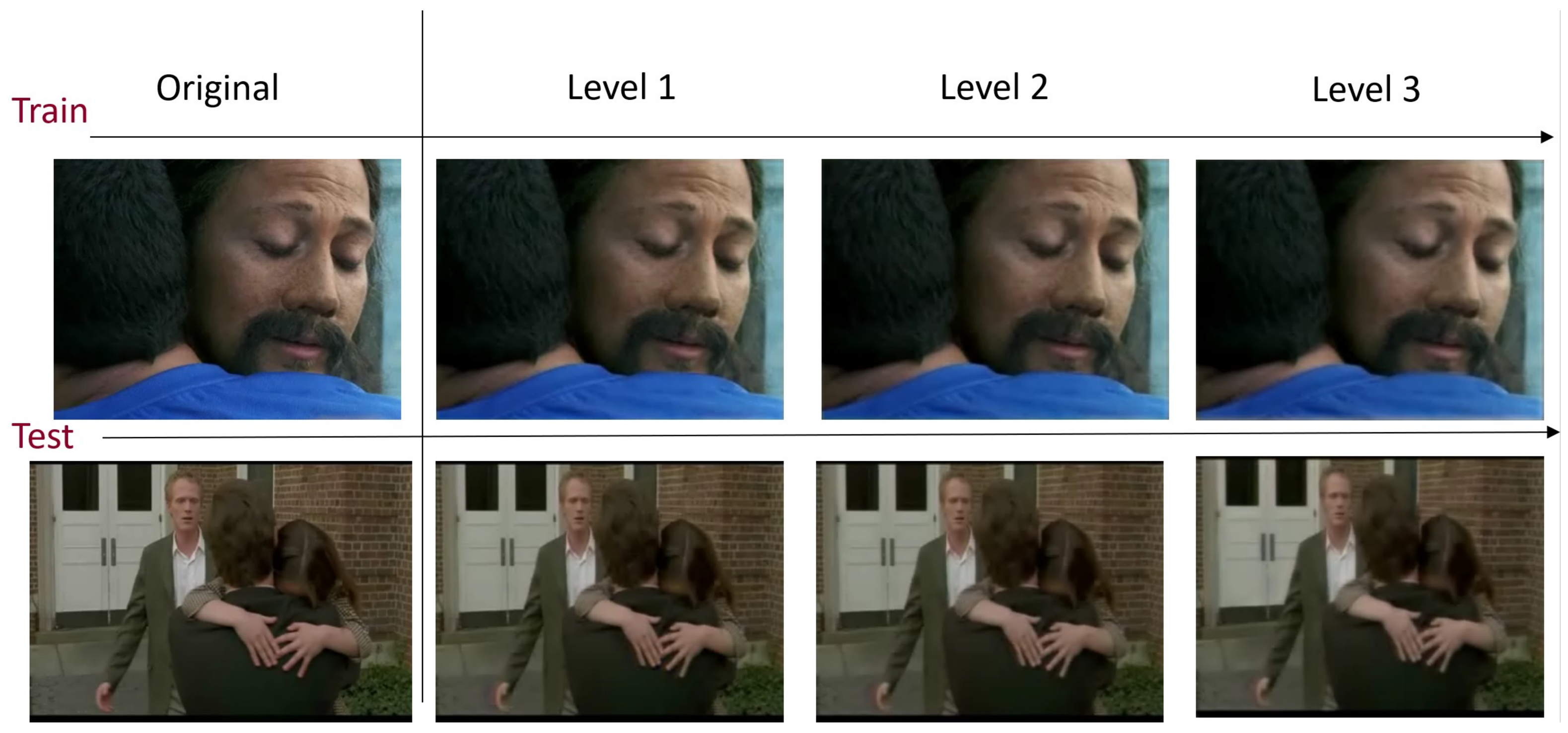
- **Layered Block Development:** Implement dynamic autoencoder with layered blocks for multiple compression paths and real-time compression ratio adjustment.
- **Multi-Level Compression:** Design a 3-level compression model for images and videos, expandable to N levels, to cater to various data types.
- **Dynamic Resource Optimization:** Facilitate resource identification and optimization via metadata exchange, adjusting compression ratios based on device performance and network bandwidth.




## Results

- **Example of 3-level Dynamic Neural Network:** Minor quality loss in small videos; larger media handles multiple compressions well.




<PSNR>

| | Original | Level 1 | Level 2 | Level 3 |
|---|---|---|---|---|
| Train | ∞ | 37.492 | 33.99555 | 32.0957 |
| Test | ∞ | 37.482 | 33.97180 | 32.0744 |

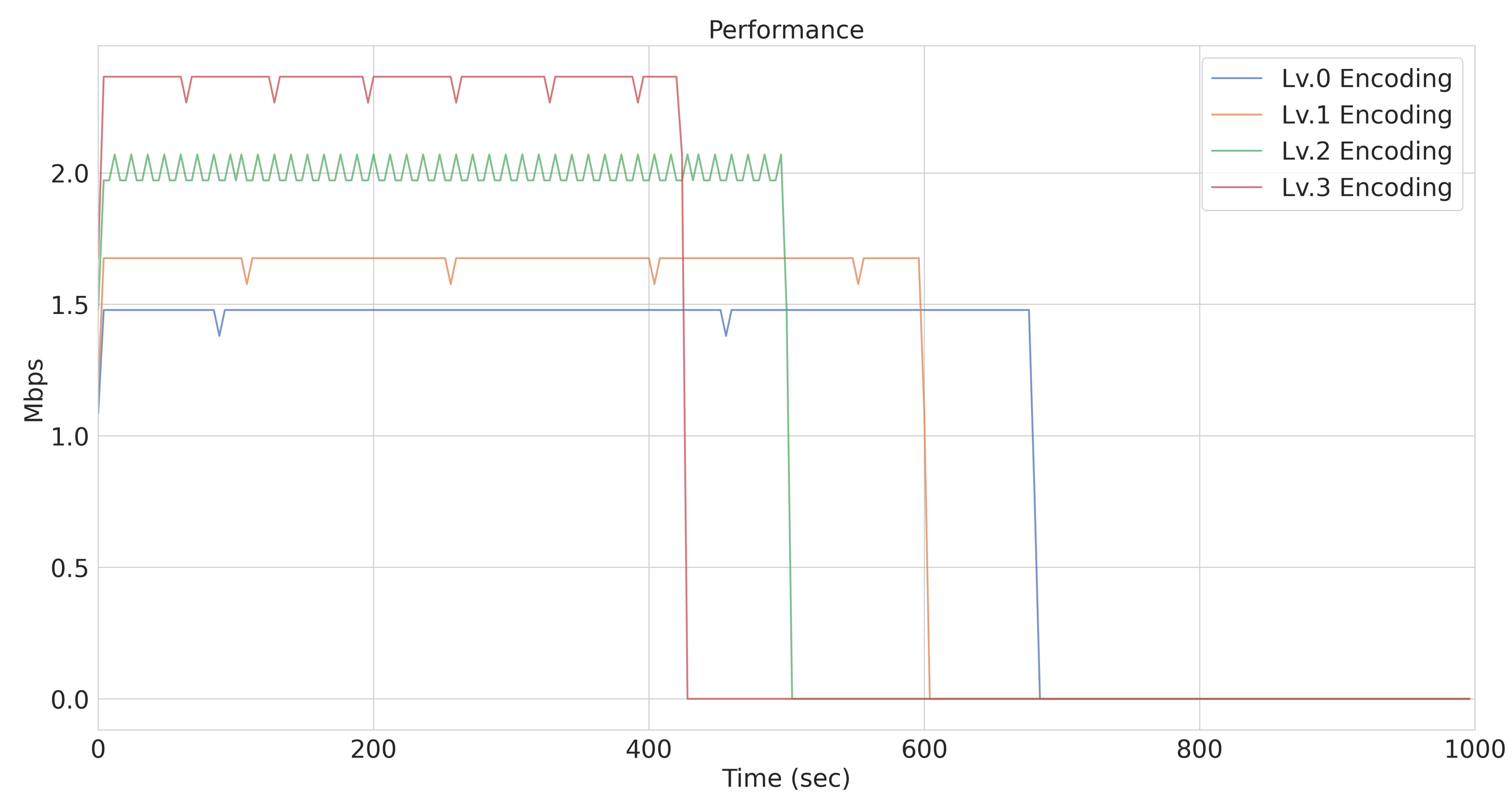
<SSIM>

| | Original | Level 1 | Level 2 | Level 3 |
|---|---|---|---|---|
| Train | 1 | 0.970872 | 0.959086 | 0.9338 |
| Test | 1 | 0.971050 | 0.959276 | 0.9341 |

- **Computational Time:**

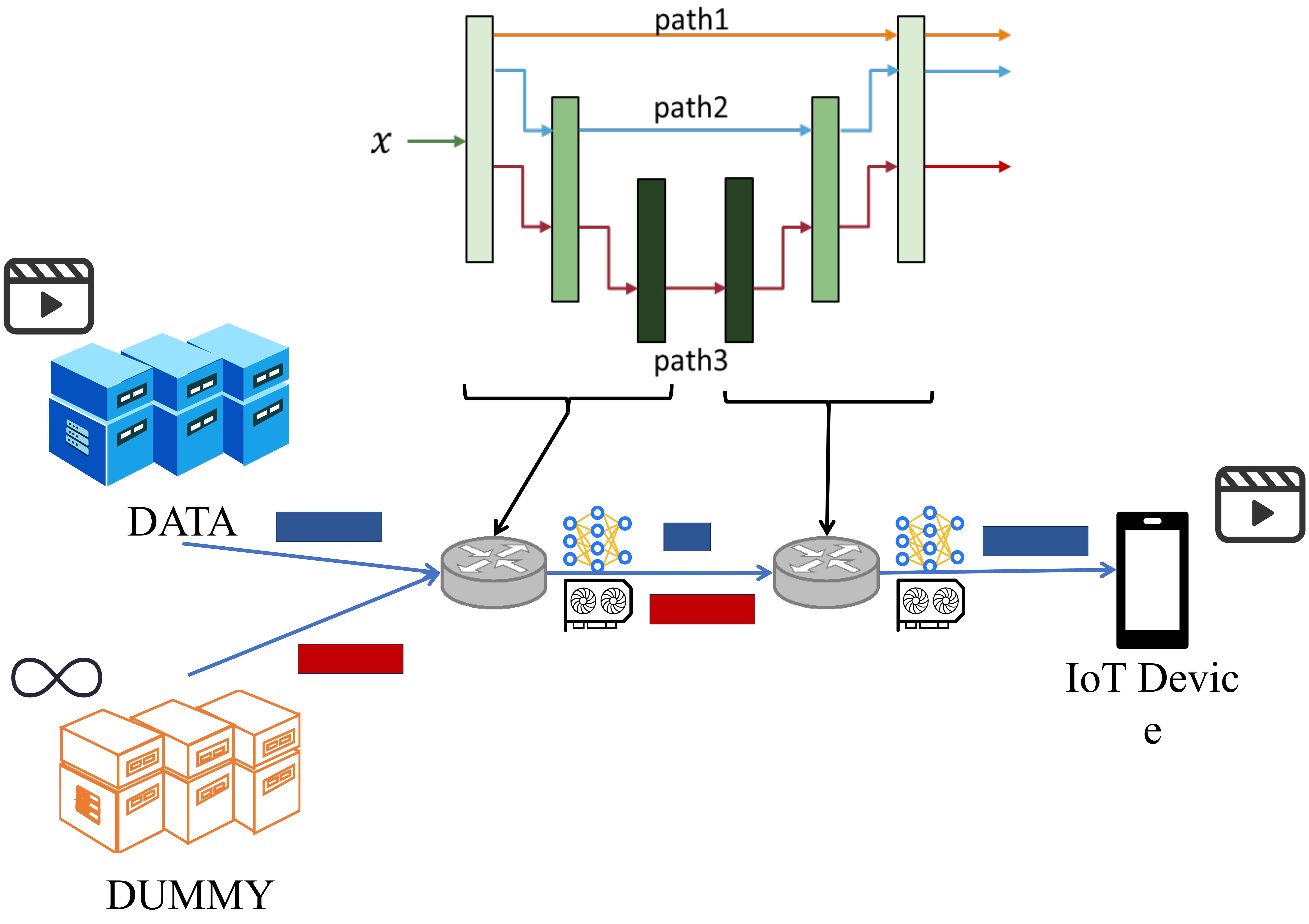
| Stage | Time Range (ms) | Average Time (ms) | Median (ms) | Variance |
|---|---|---|---|---|
| Lv1_encoder | 0.54 - 2.80 | 0.67 | 0.57 | 0.1757 |
| Lv2_encoder | 0.67 - 3.19 | 0.92 | 0.71 | 0.3640 |
| Lv3_encoder | 0.83 - 4.32 | 1.87 | 1.12 | 1.3835 |
| Lv3_decoder | 0.81 - 4.19 | 1.43 | 1.05 | 0.7343 |
| Lv2_decoder | 0.66 - 3.64 | 0.93 | 0.70 | 0.4728 |
| Lv1_decoder | 0.53 - 2.96 | 0.67 | 0.56 | 0.2102 |

- **Network Performance:** Enhanced video transmission with adjustable autoencoder computation. (The study transmits **video** and **dummy packets** of **equal size** at a frequency of **every 0.1 seconds**, over a link with a bandwidth of **3 Mbps**.)




## Conclusion

Our study presents a dynamic autoencoder for efficient multimedia compression, optimizing network and computational resources. It combines adaptive structures and intelligent resource management to enhance network performance, demonstrating a scalable solution for handling increased multimedia traffic with minimal quality loss.


path1
path2
$x$
path3
DATA
DUMMY
IoT Devic
e