RDI-Net: Relational Dynamic Inference Networks
Supplementary Material

In this Supplementary Material, we provide several detailed experiments and illustrations. First, we provide more ablation studies as discussed in Section 4.3 of the manuscript. Second, we provide detailed results of our method on CIFAR-10/100 and ImageNet as discussed in Section 4.2 of the manuscript. Finally, we provide the concrete number of the comparisons in Figure 6 of the manuscript.

A. Additional Studies

A.1. More Validation of Our Proposals

We provide more ablation studies of our proposals based on ResNet-32 and ResNet-110 in Table A. USW: Uniform Sampling Warm-up; SRM: Sample Relation Module. The first row refers to using routers without router feature fusion.

Table A. More ablations on CIFAR-10 and CIFAR-100.

Backbones	Methods			CIFAR-10		CIFAR-100	
	USW	Relational Routers	SRM	GFLOPs	Accuracy (%)	GFLOPs	Accuracy (%)
ResNet-110	—	—	—	0.30	93.21	0.29	70.32
	✓	—	—	0.28	93.52	0.27	70.81
	✓	✓	—	0.31	94.57	0.30	73.71
	✓	✓	✓	0.25	95.06	0.26	74.11
ResNet-32	—	—	—	0.09	91.55	0.10	67.42
	✓	—	—	0.09	91.92	0.11	68.76
	✓	✓	—	0.10	92.51	0.12	70.12
	✓	✓	✓	0.09	92.98	0.12	70.54

As illustrated in Table A, we present the improvement on CIFAR-10 and CIFAR-100 brought by our proposed USW (Uniform Sampling Warm-up), relational routers, and SRM (Sample Relation Module). The experiments show the effectiveness of each module on both CIFAR-10 and CIFAR-100 based on ResNet-32 and ResNet-110, respectively.

A.2. Comparison with other fusion methods.

We employ different router feature fusion methods in our relational routers, including: Average Pooling, Max Pooling, LSTM, and Graph Convolution. For LSTM, we follow the settings in SkipNet, which uses a share LSTM module for all routers. As shown in Table B, the GCN achieves a better performance in accuracy-cost trade-off.

Table B. Comparison of the router feature fusion method on CIFAR-10 with ResNet-110. GCN is used in our method.

Methods	Average Pooling	Max Pooling	LSTM	GCN
GFLOPs	0.31	0.32	0.18	0.25
Accuracy (%)	93.81	88.06	93.41	**95.06**

A.3. Methods to Measure the Distance between Samples

We employ different methods to measure the similarity of samples, including: K-Means, Structural Similarity [R1], and Histogram Similarity [R2] on CIFAR-10. For K-Means, we use an extra well-trained model for feature extraction. Differently, Structural Similarity and Histogram Similarity utilize statistical information without an extra model. Then, we introduce the Kendall rank correlation coefficient [R3], *i.e*., Kendall's τ. The Kendall's τ is computed as follows

$$\tau = \frac{P - Q}{P + Q}, \tau \in [-1, 1], \tag{i}$$

where P is the number of pairs that are concordant (in the same order in both rankings) and Q denotes the number of pairs that are discordant (in the reverse order). As shown in Table C, the best method to obtain the similarity between samples is H-Sim. When employing the unsupervised clustering method, K-Means, it only gets a result of relatively low performance compared with statistical methods. In terms of S-Sim, it achieves 94.76% performance with 0.07 Kendall rank correlation coefficient. In RDI-Net, we take H-Sim as the metric of correlation of the SRM module, which achieves the best performance and better correlation between executing paths and samples.

Table C. Comparison of different measurement methods for similarity of samples on CIFAR-10 with ResNet-110. The Kendall's τ is calculated between path similarity and the sample similarity in each batch. Note that the $\tau > 0$ can be considered as a high value, which means more than half of the rankings are consistent.

Methods	K-Means	Structural Similarity	Histogram Similarity
Kendall's τ	0.06	0.07	0.11
Accuracy (%)	90.71	94.76	95.06

[R1] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. In *IEEE Trans. Image Process.*, 2004.

[R2] Rafael C Gonzales and Richard E Woods. In *Digital image processing.*, 2002.

[R3] Maurice G Kendall. A new measure of rank correlation. *957 Biometrika, 30(1/2):81–93.*, 1938.

B. Detailed Results of RDI-Net under Different Executing Rates

As discussed in Fig.5 in the manuscript, the detailed results of RDI-Net on CIFAR-10/100 and ImageNet under different executing rates t is shown in Table D, Table E, and Table F.

B.1. CIFAR-10

Table D. Results on CIFAR-10 based on ResNets-110. The Rate refers to the target executing rate (t).

CIFAR-10			
Methods	Rates	Accuracy (%)	GFLOPs
ResNet-110	—	93.60	0.50
RDI-Net110	0.4	92.98	0.09
RDI-Net110	0.5	93.97	0.16
RDI-Net110	0.6	95.06	0.25
RDI-Net110	0.8	95.12	0.39

B.2. CIFAR-100

Table E. Results on CIFAR-100 based on ResNets-110. The Rate refers to the target executing rate (t).

CIFAR-100			
Methods	Rates	Accuracy (%)	GFLOPs
ResNet-110	—	71.20	0.50
RDI-Net110	0.2	71.22	0.11
RDI-Net110	0.4	73.23	0.20
RDI-Net110	0.6	74.04	0.26
RDI-Net110	0.8	74.30	0.36

B.3. ImageNet

Table F. The performance of our method on ImageNet. The Rate refers to the target executing rate (t).

ImageNet				
Methods	Backbones	Rates	GFLOPs	Accuracy (%)
ResNet-50	—	—	7.72	75.30
RDI-Net50	ResNet-50	0.6	4.96	75.28
RDI-Net50	ResNet-50	0.7	6.02	76.47
RDI-Net50	ResNet-50	0.8	6.68	76.89
RDI-Net50	ResNet-50	0.9	7.60	77.01
ResNet-101	—	—	15.26	76.40
RDI-Net101	ResNet-101	0.5	8.12	77.27
RDI-Net101	ResNet-101	0.7	11.14	77.68

C. Detailed Results in Performance Comparison

C.1. CIFAR-10

Table G. The The concrete number of performance comparison on CIFAR-10 as discussed in Fig.6.a in the manuscript.

Comparison with SOTAs on CIFAR-10			
Methods	Backbone	GFLOPs	Accuracy (%)
Baselines			
ResNet-110	—	0.50	93.60
ResNet-32	—	0.14	92.40
Dynamic inference			
SkipNet	ResNet-110	0.09	92.38
BlockDrop	ResNet-110	0.17	93.60
Conv-AIG	ResNet-110	0.41	94.24
CoDiNet	ResNet-110	0.29	94.47
CoDiNet (light)	ResNet-110	0.10	92.45
Early prediction			
IamNN	ResNet-101	1.10	94.60
DG-Net	ResNet-101	3.20	93.99
DG-Net (light)	ResNet-101	2.22	91.99
DDI	ResNet-74	0.14	93.88
RDI-Net (light)	ResNet-32	0.09	92.98
RDI-Net	ResNet-110	0.25	**95.06**

C.2. ImageNet

Table H. The concrete number of performance comparison on ImageNet as discussed in Fig.6.c in the manuscript.

Comparison with SOTAs on ImageNet			
Methods	Backbone	GFLOPs	Accuracy (%)
Baselines			
ResNet-50	——	7.72	75.30
ResNet-101	——	15.26	76.40
Dynamic inference			
Conv-AIG	ResNet-50	6.12	76.18
SkipNet	ResNet-101	7.22	75.22
CoDiNet	ResNet-50	6.24	76.43
DG-Net	ResNet-101	4.40	74.70
DR-ResNet	ResNets	6.24	76.48
Early prediction			
DDI	DenseNet-201	7.00	76.50
MSDN	DenseNets	4.60	74.24
RA-Net	DenseNets	4.80	75.10
Ours			
RDI-Net50	ResNet-50	6.02	76.47
RDI-Net101	ResNet-101	11.14	**77.68**