Supplementary Material

A Lightweight Network Architecture

As shown in Figure 13, the fully connected neural network is used to train neural-based differential-linear distinguishers. The input sample is reshaped into a $1 \times t$-matrix. Subsequently, this matrix is subjected to sequential processing by two dense layers with t units. The final output is generated by a `Dense` layer with sigmoid activation.

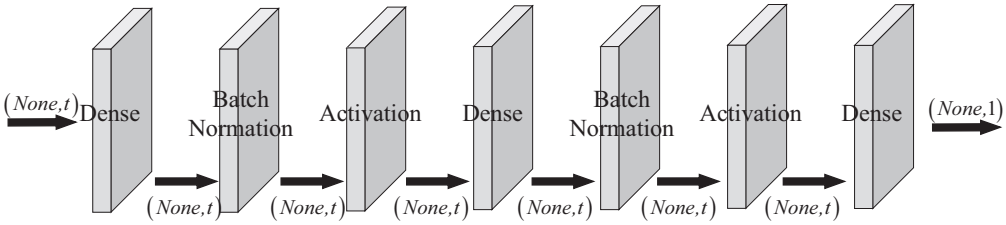

Fig. 13: Lightweight Network Architecture

In contrast to the network shown in Figure 4, the lightweight network eliminates the use of both convolutional layers and residual towers.

Employing the network architecture shown in Figure 13, we train the differential-linear distinguishers for 7-round DES. We employ the same parameters – including batch size, optimizer, learning rate, and number of epochs – as those used for training the distinguishers described in Section 4.1. The new distinguishers are shown in Table 6.

Table 6: Summary of accuracy of $ND_7{}^t$ using the lightweight network. TPR: True Positive Rate; TNR: True Negative Rate; FPR: False Positive Rate; FNR: False Negative Rate.

$ND_7{}^t$	*Accuracy*	TPR	TNR	FPR	FNR	FLOPs	Parameter Counts
$ND_7{}^{32}$	54.18%	51.53%	56.82%	43.18%	48.47%	4.609 k	2.401 k
$ND_7{}^{48}$	55.13%	54.64%	55.62%	44.38%	45.36%	9.985 k	5.137 k
$ND_7{}^{64}$	55.92%	56.70%	55.15%	44.85%	43.30%	17.409 k	8.897 k
$ND_7{}^{80}$	56.60%	58.43%	54.78%	45.22%	41.57%	26.881 k	13.681 k
$ND_7{}^{96}$	57.21%	59.79%	54.63%	45.37%	40.21%	38.401 k	19.489 k
$ND_7{}^{112}$	57.80%	55.36%	60.24%	39.76%	44.64%	51.969 k	26.321 k
$ND_7{}^{128}$	58.36%	57.08%	59.64%	40.36%	42.92%	67.585 k	34.177 k

To assess the performance of the distinguisher resulting from training the new architecture relative to established distinguishers, we utilize accuracy and model complexity as comparative metrics. As shown in Table 2 and Table 6, compared to existing distinguishers, those trained with the new network architecture demonstrate reduced accuracy. On the other hand, due to less network parameters, the new architecture exhibits superior computational efficiency compared to existing distinguishers.