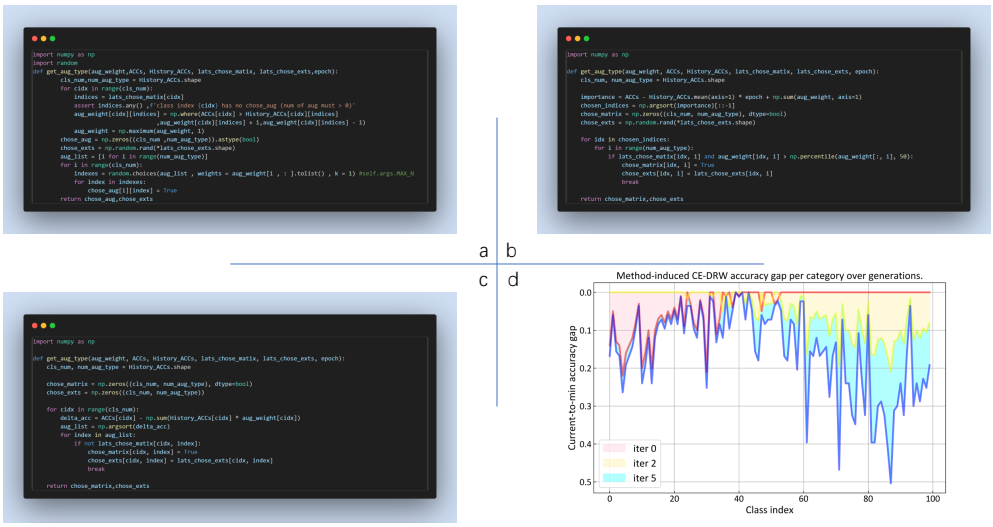

Figure 1: Demonstration of some generated methods (a, b, c) and the performance changes during the evolution process (d).

Table 1: Accuracy (%) on CIFAR-100-LT dataset.

Method	IR = 50				IR = 100			
	Head	Medium	Tail	All	Head	Medium	Tail	All
CE-DRW	60.6	39.0	22.9	45.0 (+0.0)	63.4	41.2	15.7	41.4 (+0.0)
CE-DRW + CMO[2]	60.8	48.1	39.1	51.9 (+6.9)	56.3	45.4	34.2	46.0 (+4.6)
CE-DRW + H2T[1]	63.3	55.9	34.5	52.4 (+7.4)	61.0	51.2	25.3	47.2 (+5.8)
CE-DRW + SimpleLLM	62.3	49.4	38.8	52.7 (+7.7)	62.1	49.6	27.9	47.5 (+6.1)
CE-DRW + LLM-AutoDA	63.1	48.4	39.3	52.8 (+7.8)	62.9	50.7	29.9	48.7 (+7.3)
BS	60.3	41.3	34.3	47.9 (+0.0)	59.6	42.3	23.7	42.8 (+0.0)
BS + CMO	64.0	46.5	27.4	50.4 (+2.5)	63.1	48.8	26.7	47.2 (+4.4)
BS + H2T	65.2	54.0	35.8	52.8 (+4.9)	63.5	50.0	26.6	48.1 (+5.3)
BS + SimpleLLM	62.4	48.8	37.7	52.4 (+4.5)	62.4	48.8	30.6	48.1 (+5.3)
BS + LLM-AutoDA	63.3	50.5	40.2	53.9 (+6.0)	63.3	50.0	31.0	49.0 (+6.2)

References

[1] Mengke Li, HU Zhikai, Yang Lu, Weichao Lan, Yiu-ming Cheung, and Hui Huang. Feature fusion from head to tail for long-tailed visual recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 13581–13589, 2024.

[2] Seulki Park, Youngkyu Hong, Byeongho Heo, Sangdoo Yun, and Jin Young Choi. The majority can help the minority: Context-rich minority oversampling for long-tailed classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2022.