Table 6: The hyperparameter setting of our proposed load recognition approach and the compared models.

Hyperparameter	Vit	Cait	ConvNeXt	Swin	ResSENet
learning rate	1e-4	1e-4	1e-4	1e-4	8e-4
optimizer	Adam	Adam	Adam	Adam	AdamW
batch size	16	16	16	16	32
patch size	4	4	-	8	-
layers	6	(6, 2)	(3, 3, 9, 3)	(2, 2, 6, 2)	(1, 1, 3, 1)
heads	8	8	-	(3, 6, 12, 24)	-