Evaluation of continuous image features learned by ODE Nets

Fabio Carrara, Giuseppe Amato, Fabrizio Falchi, Claudio Gennaro

fabio.carrara@isti.cnr.it

GitHub

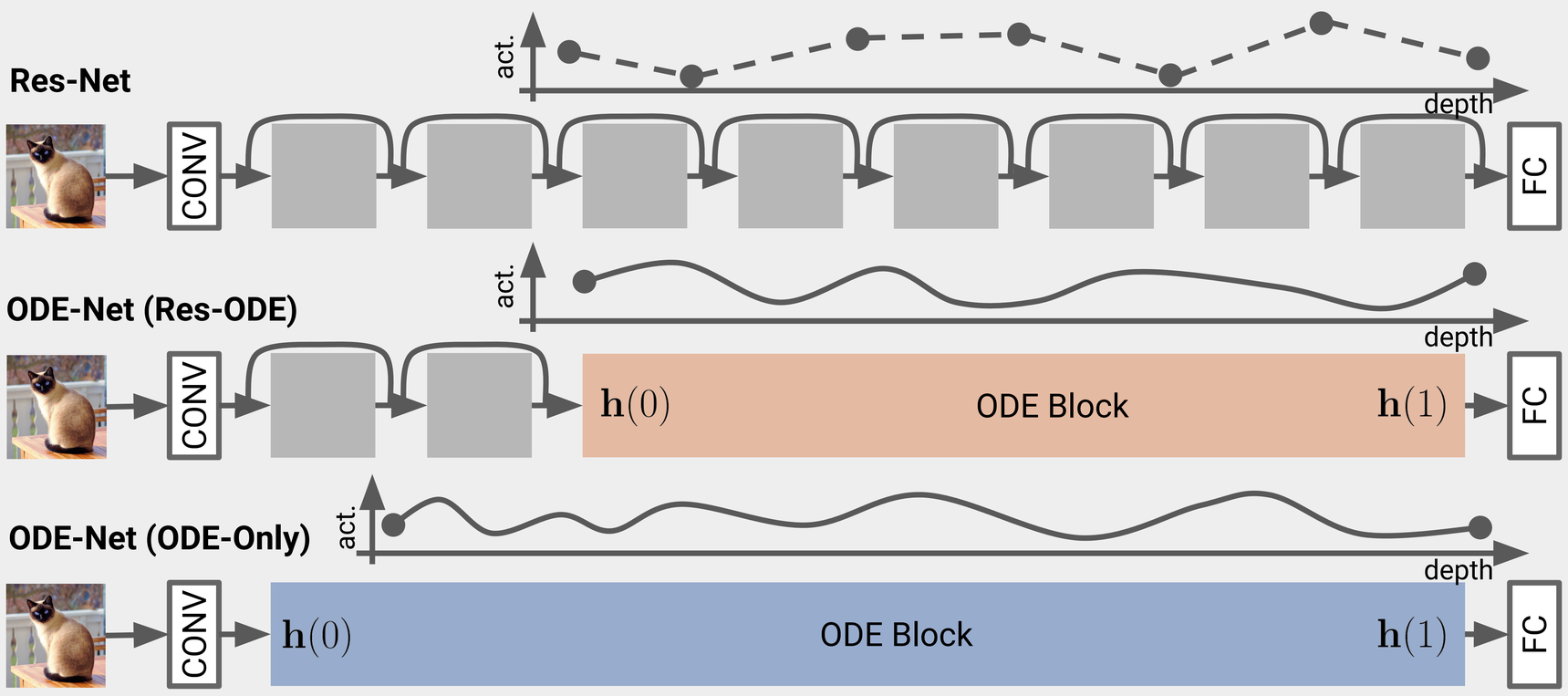
ODE-Nets [1]

- net defined by ordinary differential equations (ODE)
- forward/backward computed by generic ODE solver
- trained with standard gradient-based optimizers

[1] Chen, Tian Qi, et al. "Neural ordinary differential equations." Advances in neural information processing systems. 2018.

Properties of ODE Nets

- reversibility → ***O(1) memory consumption***
- adaptive solver → ***per-sample adaptive computation***
- adaptive solver → ***speed-accuracy trade-off at test***
- continuity → ***natural continuous-time modeling***



Residual Block

$$\mathbf{h}_{t+1} = \mathbf{h}_t + f(\mathbf{h}_t, \theta_t)$$

ODE Block

$$\frac{\mathrm{d}\mathbf{h}(t)}{\mathrm{d}t} = f(\mathbf{h}(t), t, \theta)$$

Invoke ODE solver to find solution:

$$\mathbf{h}(1) = \mathbf{h}(0) + \int_0^1 f(\mathbf{h}(t), t, \theta)\,\mathrm{d}t$$

Continuous Image Features Evaluation

- image classifiers with ODE blocks
- continuously evolving feature representation $\mathbf{h}(t)$
- sample (potentially infinite) features with $t \in [0, 1]$
- evaluate features in transfer learning task

Experiments & Results

- Train feature extractors on CIFAR-10 dataset
- Evaluate on Tiny-IN-200 classification (SVM + feats.)
- ➕ **less parameters, better transferability**
- ➖ currently slow (slower as training proceeds)

Model	CIFAR-10 Test Error	Model Parameters	ODE Solver Steps
Res-Net	7.28%	7.92M	-
Res-ODE	7.80%	2.02M	3.8 ± 0.4
ODE-only (our)	9.17%	1.20M	7.8 ± 1.5