Deep Scattering Network with Max-pooling

Taekyung Ki[a] and Youngmi Hur[*,1,a,b]

[a]School of Mathematics and Computing (Mathematics), Yonsei University, South Korea
[b]School of Mathematics, Korea Institute for Advanced Study, South Korea
[*]Corresponding author: yhur@yonsei.ac.kr

Scattering network is a convolutional network, consisting of cascading convolutions using pre-defined wavelets followed by the modulus operator [1]. The scattering network is one of few mathematical tools explaining the convolutional neural networks (CNNs). However, a pooling operator, which is a main component of CNNs, is not considered in the original scattering network. We model a continuous max-pooling, apply it to the scattering network, and get a new network named **scattering-maxp network**. We show that the scattering-maxp network shares many useful properties of the scattering network including translation invariance, and we conduct numerical experiments showing the computational advantage of our network. We summarize our main findings below. More details about our results can be found in [2] and Python codes are available in `https://github.com/TaekyungKi/Scattering_maxp`.

Let $f \in L^\infty(\mathbb{R}^d)$ be supported in a compact rectangular region $D = \cup_{i=1}^{N} D^{(i)} \subset \mathbb{R}^d$. The continuous max-pooling P is defined by $\mathrm{P}(f) := \sum_i \|f\chi_{D^{(i)}}\|_\infty \chi_{D^{(i)}}(S\cdot)$. Using P and a low-pass kernel ϕ_{2^J}, we define an operator $\tilde{S}_J[q]f := \phi_{2^J} * \tilde{U}[\lambda_m] \cdots \tilde{U}[\lambda_2]\tilde{U}[\lambda_1]f$, where $q := (\lambda_1, \lambda_2, \cdots, \lambda_m)$ and $\tilde{U}[\lambda_k]f := \mathrm{P}(|\psi_{\lambda_k} * f|)$ with the wavelet kernel ψ_{λ_k}. The outputs $\{\tilde{S}_J[q]f\}_{q\in\Lambda_J^m}$ constitute the m-th layer of the scattering-maxp network.

Result 1. Suppose that $c \in \mathbb{R}^d$ satisfies $0 \in D + c$ and that $|\hat{\phi}(\omega)||\omega| < B$, a.e. $\omega \in \mathbb{R}^d$, for some $B > 0$. For the translation operator $T_c f(x) := f(x - c)$, we have

$$\lim_{m\to\infty} \sum_{q\in\Lambda_J^m} \|\tilde{S}_J[q]f - \tilde{S}_J[q](T_c f)\|_2^2 = 0.$$

Result 2. Image classification results of Caltech-101 (resp. Caltech-256) are given in the following table. It shows the scattering-maxp achieves competitive performance and much faster in training time, when compared with the original scattering.

Model	**# of parameters** [m]	**Accuracy** [%]	**Training time** [s/epochs]
Scattering	87.5 (resp. 87.6)	98.49 (resp. 96.06)	284 (resp. 1006)
Scattering-maxp	9.9 (resp. 9.9)	98.59 (resp. 92.11)	206 (resp. 651)

[1] S. Mallat, "Group invariant scattering," *Comm. Pure Appl. Math.*, 2012.

[2] Taekyung Ki and Youngmi Hur, "Deep scattering network with max-pooling," preprint (`https://arxiv.org/abs/2101.02321`), 2021.

[1]Supported by the National Research Foundation of Korea [Grant Number 2015R1A5A1009350] in part, and by the 'Ministry of Science and ICT' and NIPA via "HPC Support" Project in part.