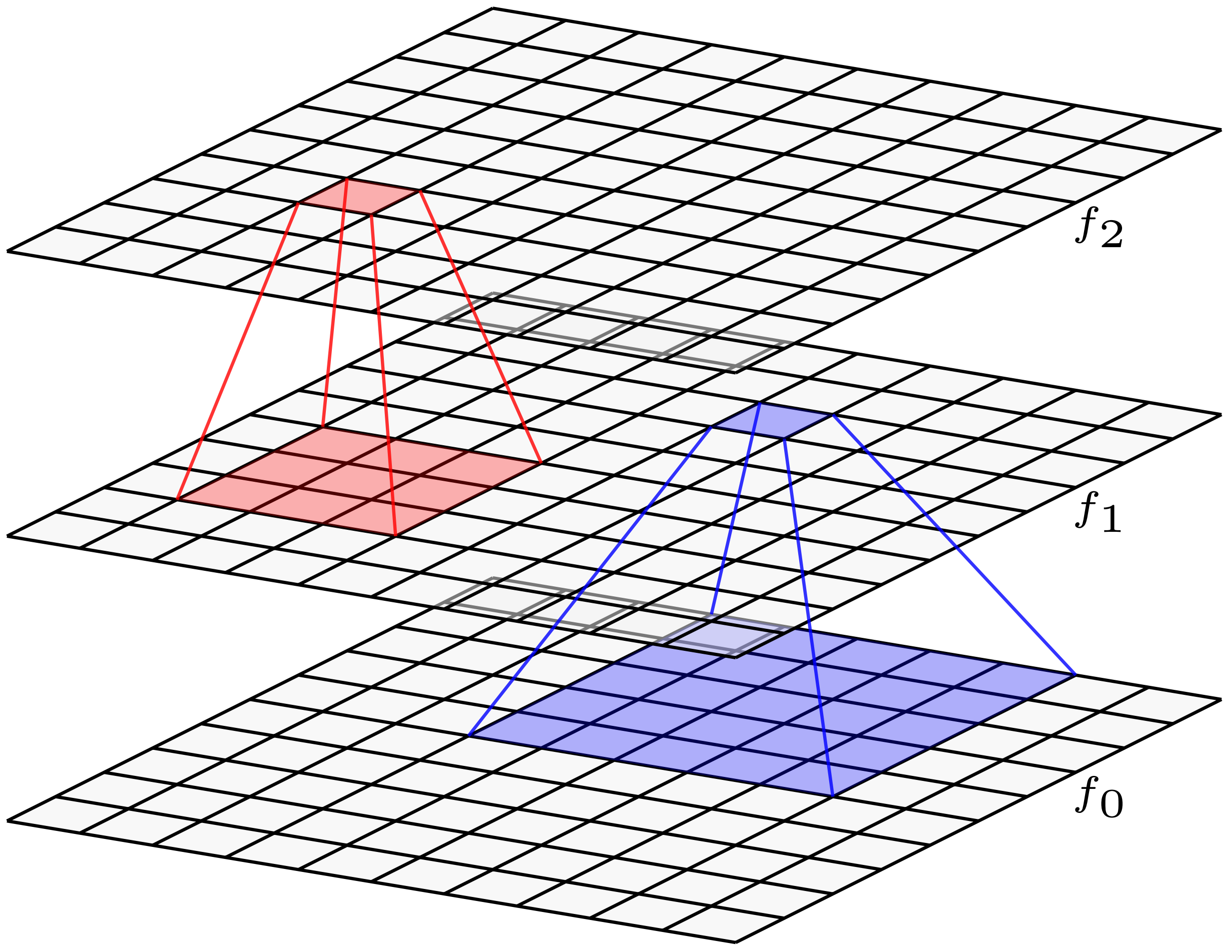
On the Generalization of Convolution to the Action of Compact Groups

Risi Kondor
The University of Chicago

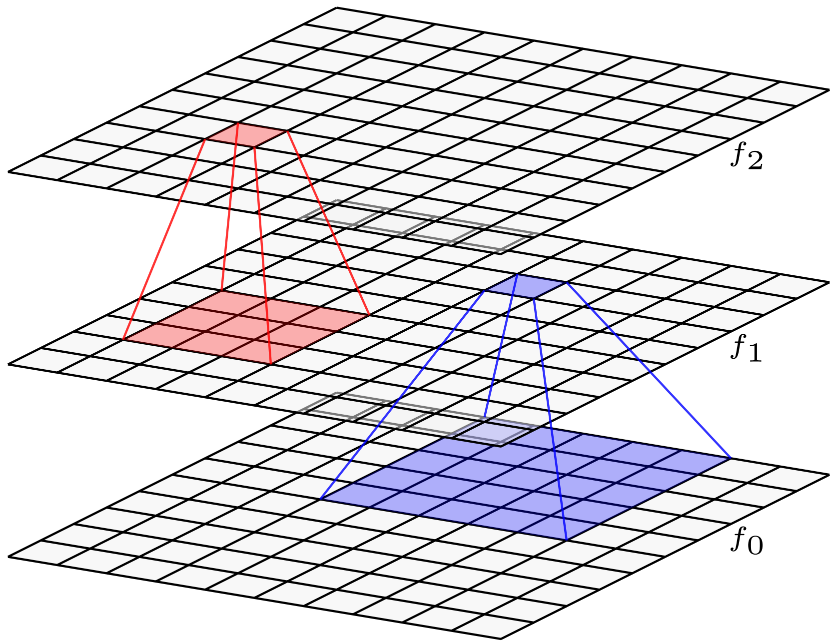
Shubhendu Trivedi
Toyota Technological Institute, Chicago

Convolutional Neural Networks



$$f_0 \longmapsto \phi_1(f_0) \xmapsto{\sigma} f_1 \longmapsto \phi_2(f_1) \xmapsto{\sigma} f_2 \longmapsto \ldots \longmapsto \phi_L(f_{\ell-1}) \longmapsto f_L$$

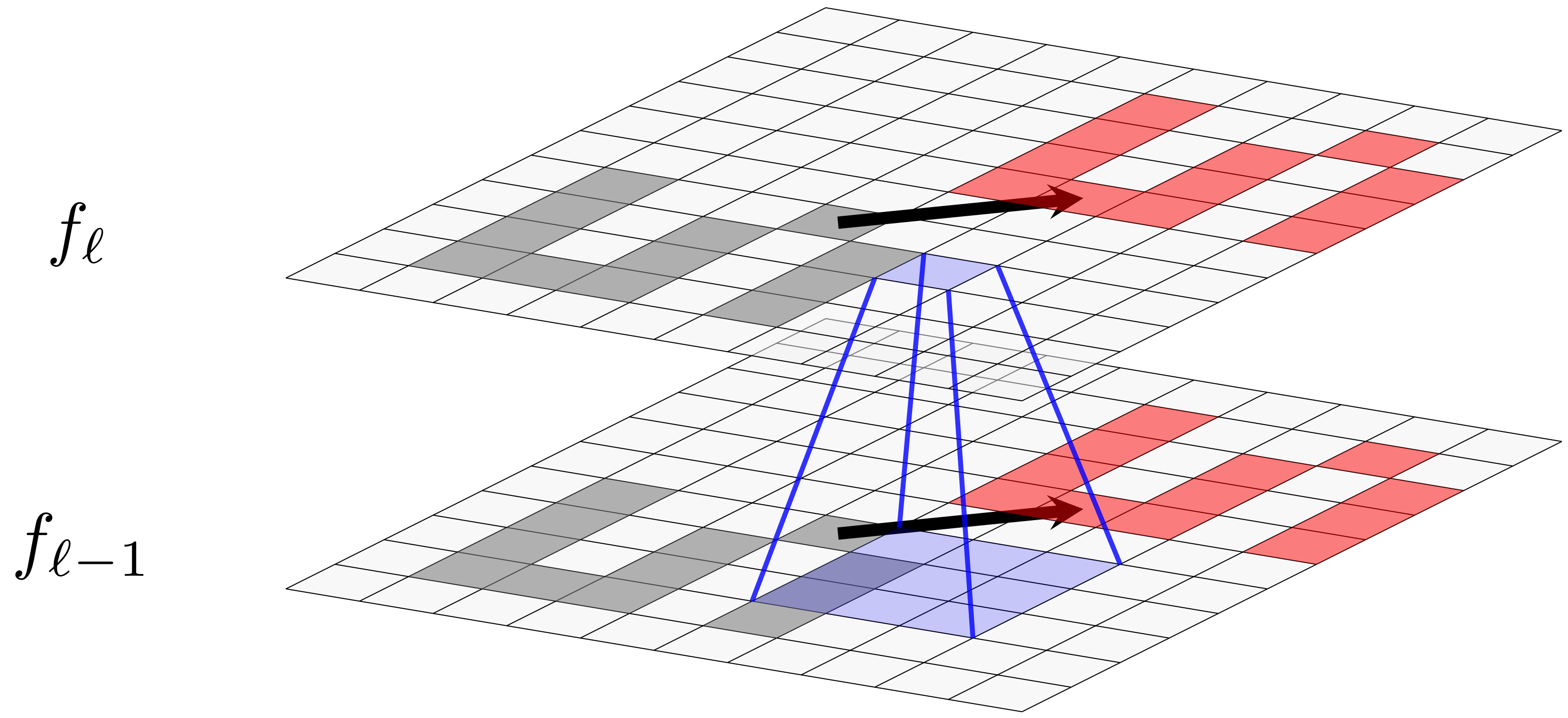
Convolutional Neural Networks



$$\phi_\ell(f_{\ell-1}) = (f_{\ell-1} * g_\ell)(x) = \sum_{y \in \mathbb{Z}^2} f_{\ell-1}(x-y)\, g_\ell(y)$$

Filter at layer ℓ

Equivariance

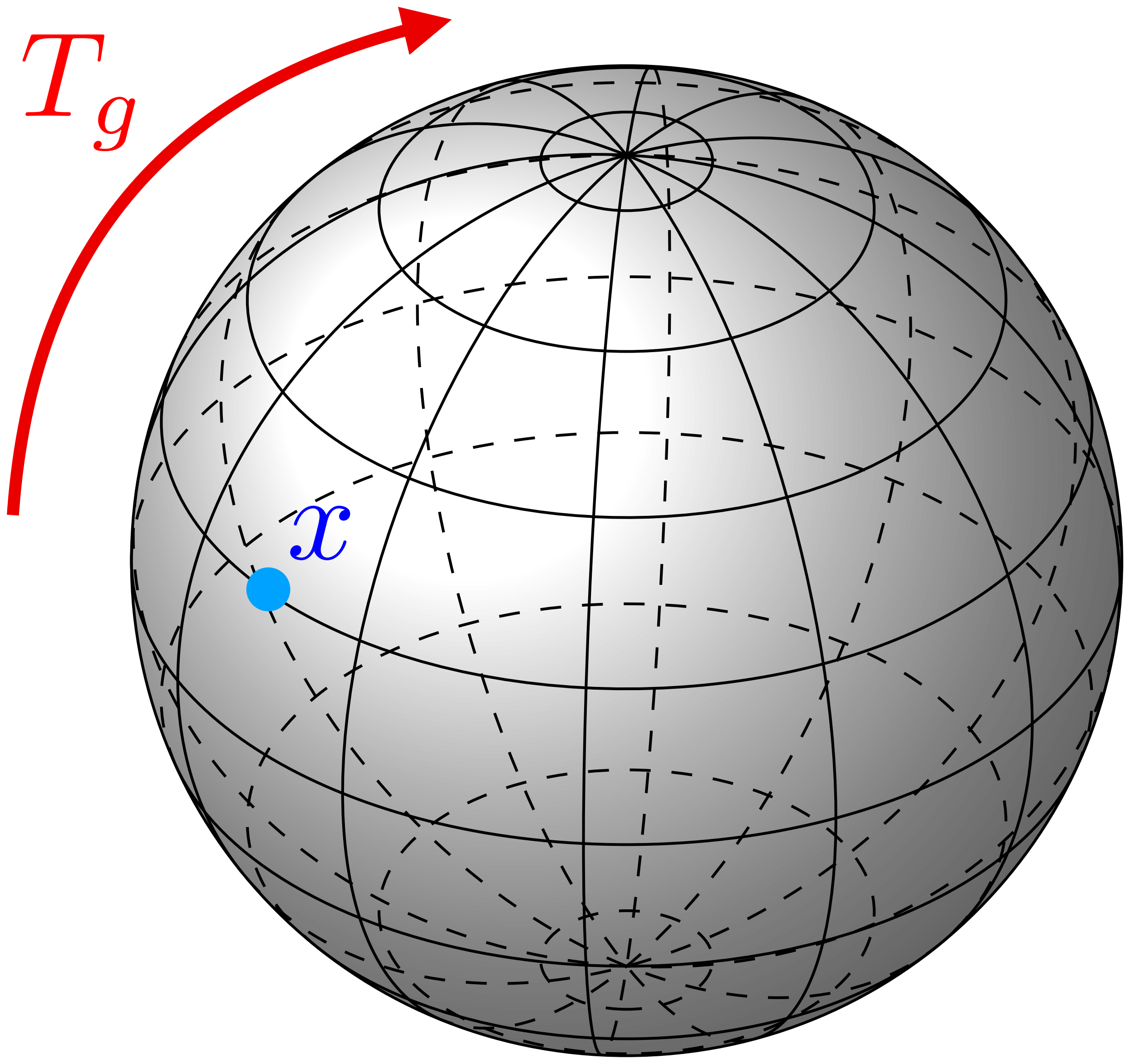


$$f'_{\ell-1}(\mathbf{x}) = f_{\ell-1}(\mathbf{x} - \boldsymbol{t}) \qquad \Longrightarrow \quad f'_\ell(\mathbf{x}) = f_\ell(\mathbf{x} - \boldsymbol{t})$$

1. Parameter sharing
2. Same filters applied to every part of the image
3. Invariance, if followed by final invariant layer.

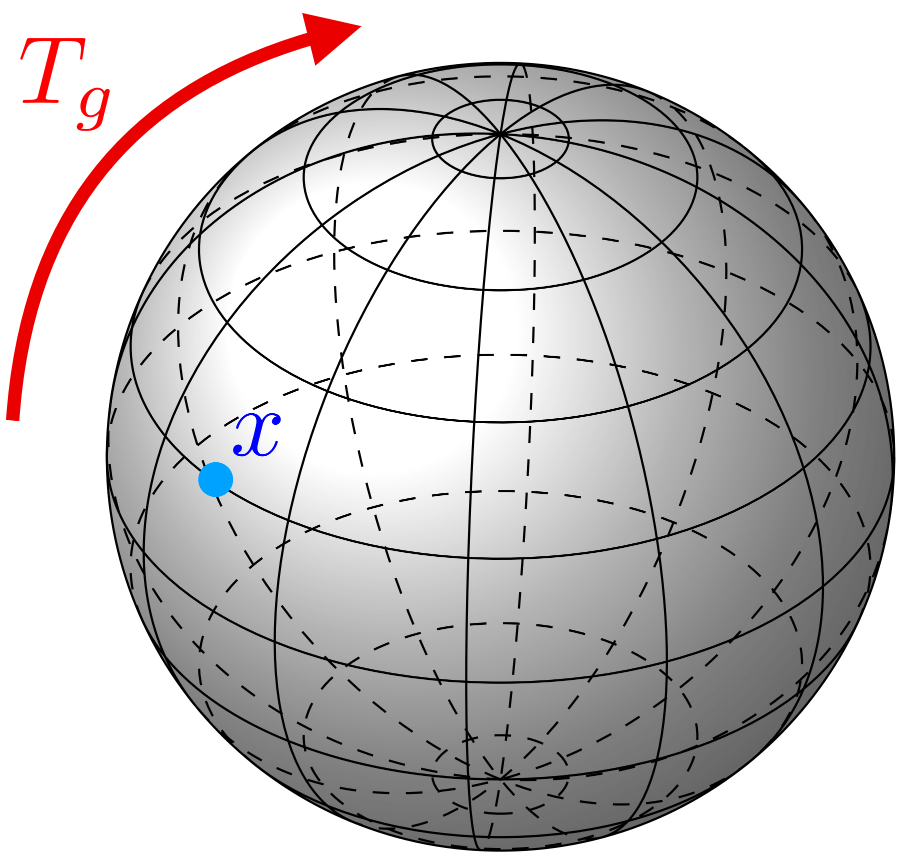
[Cohen & Welling, 2016]
[Cohen & Welling, 2017]
[Ravanbakhsh, Schneider & Poczos, 2017]

How do we generalize this to other transformation groups?



[Cohen, Geiger, Köhler & Welling, 2018]

Group actions



1. Our function lives on a space $\mathcal{X}$

$$f : \mathcal{X} \to \mathbb{C}$$

2. We have a group G acting on $\mathcal{X}$

$$x \mapsto T_g(x)$$

3. This induces an action on functions

$$f \xmapsto{\mathbb{T}_g} f' \qquad f'(x) = f(T_g^{-1}(x))$$

Equivariance

Equivariance

1. We have to different spaces $\mathcal{X}_1$ and $\mathcal{X}_2$ on which G acts by

$$T_g^{(1)}: \mathcal{X}_1 \to \mathcal{X}_1 \qquad\qquad T_g^{(2)}: \mathcal{X}_2 \to \mathcal{X}_2$$

2. We have corresponding actions on functions

$$f \mapsto \mathbb{T}_g^{(1)}(f) \qquad \mathbb{T}_g^{(1)}(f)(x) = f((T_g^{(1)})^{-1}(x))$$
$$f \mapsto \mathbb{T}_g^{(2)}(f) \qquad \mathbb{T}_g^{(2)}(f)(x) = f((T_g^{(2)})^{-1}(x))$$

3. A map $\phi: L(\mathcal{X}_1) \to L(\mathcal{X}_2)$ is equivariant to these actions if

$$\phi(\mathbb{T}_g^{(1)}(f)) = \mathbb{T}_g^{(2)}(\phi(f))$$

for all $f \in L(\mathcal{X}_1)$.

$$
\begin{tikzcd}
L(\mathcal{X}_1) \arrow[r, "\mathbb{T}_g^{(1)}"] \arrow[d, "\phi"] & L(\mathcal{X}_1) \arrow[d, "\phi"] \\
L(\mathcal{X}_2) \arrow[r, "\mathbb{T}_g^{(2)}"] & L(\mathcal{X}_2)
\end{tikzcd}
$$

$$(f * g)(u) = \int_G f(uv^{-1})\, g(v)\, d\mu(v)$$

$$(f * g)(u) = \int_G f{\uparrow}^G(uv^{-1})\, g{\uparrow}^G(v)\, d\mu(v)$$

Main theorem

A feed-forward neural network is equivariant to the action of a compact group G if and only if the linear operation in each layer is of the form

$$\phi_\ell(f_{\ell-1}) = f_{\ell-1} * g_\ell.$$

$$L(\mathcal{X}) = \overset{\rho_0}{\overset{\downarrow}{V_0}} \oplus \overset{\rho_1}{\overset{\downarrow}{V_1}} \oplus \overset{\rho_2}{\overset{\downarrow}{V_2}} \oplus \ldots \oplus \overset{\rho_p}{\overset{\downarrow}{V_p}}$$

$$V_i = W_i^1 \oplus W_i^2 \oplus \ldots \oplus W_i^{m_i}$$

Consequences

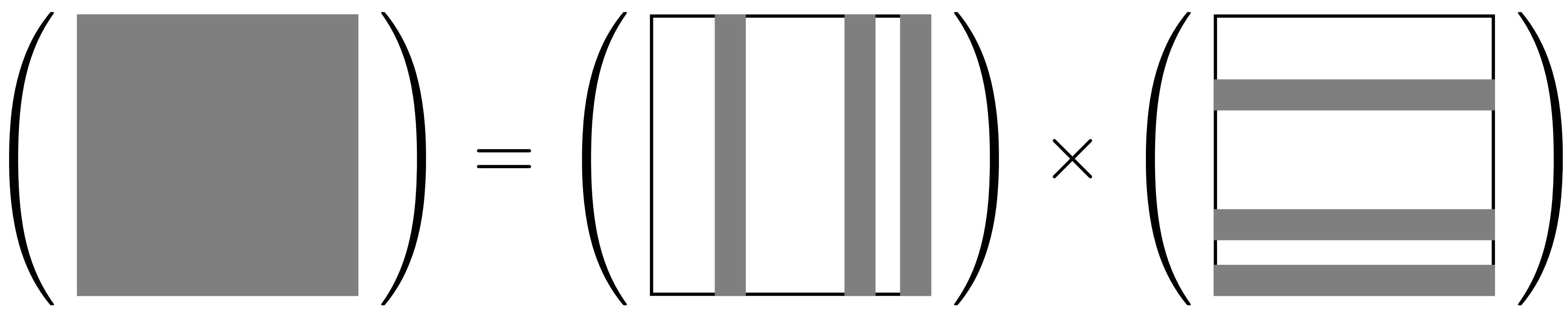
$$\widehat{f}(\rho_i) = \int_G f(u)\, \rho_i(u)\, d\mu(u) \qquad i = 0, 1, 2, \ldots$$

$$\widehat{f * g}(\rho_i) = \widehat{f}(\rho_i) \cdot \widehat{g}(\rho_i)$$

matrix multiplication

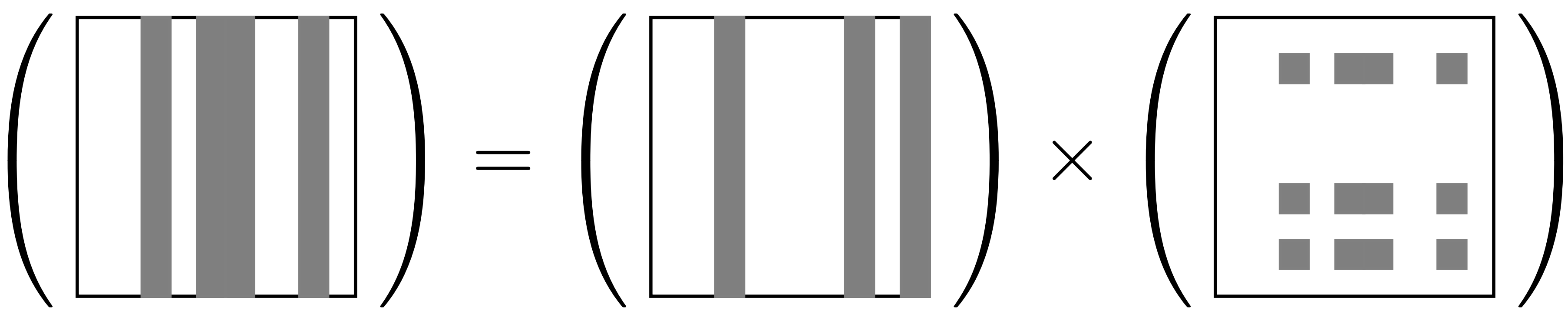
Case 1:

$f_{\ell-1}\colon G/H \to \mathbb{C}$ $\quad$ $f_\ell\colon G \to \mathbb{C}$



$\widehat{f * g}(\rho)$ $\quad$ $\widehat{f{\uparrow}^G}(\rho)$ $\quad$ $\widehat{g{\uparrow}^G}(\rho)$

Case 2: $f_{\ell-1} : G/H \to \mathbb{C}$ $f_\ell : G/K \to \mathbb{C}$



$$\widehat{f * g}(\rho) \qquad \widehat{f{\uparrow}^G}(\rho) \qquad \widehat{g{\uparrow}^G}(\rho)$$

Fourier space neural networks

1. Spherical CNNs

[Cohen, Geiger, Köhler & Welling, 2018]

[K., Lin and Trivedi, 2018]

2. Steerability and conv-nets for manifolds

[Marcos, Volpi et al., 2017]

[Masci, Boscaini et al., 2015]

[Worral, Garbin et al., 2017]

3. Neural nets for graphs

[Duvenaud at al., 2015]

[Gilmer et al., 2017]

[Son, Trivedi et al. 2018]

4. Neural nets for physical systems

[…]

Also see:

[Cohen, Geiger & Weiler, 2018]

Conclusions

1. There is a clear prescription for how to generalize CNNs to architectures that are equivariant to the action of any compact group.

2. When dealing with groups, the Fourier picture is much more compelling because it reduces all the fancy algebra to just matrix multiplication.

Poster: 154

DARPA