On the Universality of Invariant Networks

Haggai Maron    Ethan Fetaya    Nimrod Segol    Yaron Lipman
Invariant tasks

- Image classification
Invariant tasks

- Image classification

- Graph/ hyper-graph classification
Invariant tasks

• Image classification

• Graph/ hyper-graph classification

• Point-cloud / set classification

• and many more...
Goal of this paper

- Invariant neural networks are a common approach for these tasks
- This paper analyzes the expressive power of invariant models
Formal definition of group action

- Let $G \leq S_n$
- $g \in G$ acts on a vector $x \in \mathbb{R}^n$ by permuting its coordinates:
  $$gx = (x_{g^{-1}(1)}, \ldots, x_{g^{-1}(n)})$$
Formal definition of group action

\[ g \in G \text{ acts on a tensor } X \in \mathbb{R}^{n^k} \text{ by permuting its coordinates in each dimension:} \]

\[ (gX)_{i_1, \ldots, i_k} \]
Formal definition of group action

- $g \in G$ acts on a tensor $X \in \mathbb{R}^{n^k}$ by permuting its coordinates in each dimension:

$$(gX)_{i_1,\ldots,i_k} = X_{g^{-1}(i_1),\ldots,g^{-1}(i_k)}$$
Invariant and equivariant functions

**Definition:** A function $f : \mathbb{R}^n \to \mathbb{R}$ is **invariant** with respect to a group $G$ if:

$$f(gx) = f(x), \quad \forall g \in G$$
Invariant and equivariant functions

**Definition:** A function \( f : \mathbb{R}^n \to \mathbb{R} \) is **invariant** with respect to a group \( G \) if:

\[
f(gx) = f(x), \quad \forall g \in G
\]

**Definition:** A function \( f : \mathbb{R}^n \to \mathbb{R}^n \) is **equivariant** with respect to a group \( G \) if:

\[
f(gx) = gf(x), \quad \forall g \in G
\]
$G$-invariant networks

Linear equivariant layers

Linear invariant + MLP
$G$-invariant networks

$x \in \mathbb{R}^n \rightarrow x \in \mathbb{R}^{n^2} \rightarrow x \in \mathbb{R}^{n^3} \rightarrow \cdots \rightarrow x \in \mathbb{R}^{n^3} \rightarrow x \in \mathbb{R}$

Linear equivariant layers

Linear invariant + MLP
Main question: How expressive are $G$-invariant networks?
How expressive are $G$-invariant networks?

Continuous Functions
(approximable with FC networks)
How expressive are $G$-invariant networks?

- Continuous Functions
  (approximable with FC networks)

- Continuous
  $G$-Invariant functions
How expressive are $G$-invariant networks?

Continuous Functions (approximable with FC networks)

$G$-Invariant functions

$G$-Invariant networks
How expressive are $G$-invariant networks?

Continuous Functions (approximable with FC networks)

Gap?

Continuous $G$-Invariant functions

$G$-Invariant networks
Theoretical results
Universality of high-order networks

**Theorem 1.** $G$-invariant networks are universal.

Tensor order might be as high as \( \binom{n}{2} \)
Lower bound on network order

**Theorem 2.** There exists groups $G \leq S_n$ for which the tensor order should be at least $O(n)$ in order to achieve universality.

Tensor order **must** be at least $\frac{n-2}{2}$. 
Theorem 3. Let $G \in S_n$. If first order $G$-invariant networks are universal, then $|[n]^2/H| < |[n]^2/G|$ for any strict super-group $G < H \leq S_n$. 

Tensor order is 1
The End

- Support
  - ERC Grant (LiftMatch)
  - Israel Science Foundation

- Thanks for listening!

“Invariant Graph Networks”
by Yaron Lipman
Saturday 11am, Grand Ballroom B
Learning and Reasoning with Graph-Structured Representations workshop