Introduction to Robust Machine Learning with Geometric Methods for Defense Applications

Pierre-Yves Lagrave, Frédéric Barbaresco

Introduction to Robust Machine Learning with Geometric Methods for Defense Applications

Pierre-Yves Lagrave[1][0000−0002−5774−636X] and Frédéric Barbaresco[2][0000−0003−3664−3609]

[1] Thales Research and Technology, Palaiseau, France
pierre-yves.lagrave@thalesgroup.com
[2] Thales Land and Air Systems, Meudon, France
frederic.barbaresco@thalesgroup.com

Abstract. This paper aims at motivating the use of geometrically informed Machine Learning algorithms for Defense applications by providing intuitions with respect to their underlying mechanisms and by shedding light on successful applications such as remote sensing imagery, radar Doppler signal processing, trajectory prediction, physical model simulation and kinematics recognition. We in particular discuss the use Equivariant Neural Networks (ENN) which achieve geometrical robustness by-design and which also appear more robust to adversarial attacks. We will also highlight how Lie Group Statistics and Machine Learning techniques can be used to process data in their native geometry. Both technologies have a wide range of applications for the Defense industry and we generally believe that exploiting the data geometry and the underlying symmetries is key to the design of efficient, reliable and robust AI-based Defense systems.

Keywords: Equivariant Neural Networks · Geometric Deep Learning · Lie Group Statistics and Machine Learning · Robustness-by-design

1 Introduction and Motivations

Conventional Deep Learning algorithms only encode limited priors about robustness to perturbations into their design. Taking the example of computer vision tasks, CNN enforce local robustness with respect to translations but they have been shown to lack of robustness with respect to other transforms such as rotations, scaling, lightening, small noise, etc. To remedy this issue, a practical approach referred to as data augmentation consists in augmenting the training set by applying transformations to the original data, typically on-the-fly during the gradient descent routine [28]. Although widely applied by practitioners because of its empirical success and implementation convenience, data augmentation is not fully satisfactory as learning invariances directly form the (augmented) data consumes significant algorithmic capacity and therefore requires models with a large number of trainable parameters, which may not be aligned with operational constraints such as memory footprint limitations or inference timing performance. Another caveat with this type of approach is that,

although recent attempts have been made with respect to the formalization of the method through group theory [7], we are still generally lacking of theoretical guarantees with respect to the behavior the algorithms trained with augmented data. Finally, data augmentation has also been shown suboptimal [12] when compared to approaches consisting in using architectures where group-based invariance/equivariance is natively enforced.

In this context, Equivariant Neural Networks (ENN), which belong to the field of Geometric Deep Learning [6], are becoming more and more popular thanks to their conceptual soundness and to their ability to reach state-of-the-art accuracies for a wide range of applications [16,13]. In particular, the underlying equivariant and/or invariant layers of ENN (see Figure 1) allow building architectures robust to generic geometrical transforms, therefore making the use of ENN a reasonable alternative to data augmentation techniques. Another way of improving the geometrical robustness of Machine Learning models is to have them operate on the native representation of the input data. In this context, Lie Group Statistics and Machine learning techniques [1,5,2,20] provide statistical approaches for data lying within Lie Groups and can therefore been seen as a particular type of learning methods on symmetric manifolds, which are compatible with the notion of equivariance as illustrated by the notion of Gauge Equivariant Neural Networks [30].

2 Encoding Symmetries with Equivariant Neural Networks

After having motivated the specification of ENN, we give in the following a high level description of their underlying mechanisms and shed light on practical applications. We refer the interested reader to [6,15] for more detailed surveys and thorough descriptions of the associated theory.

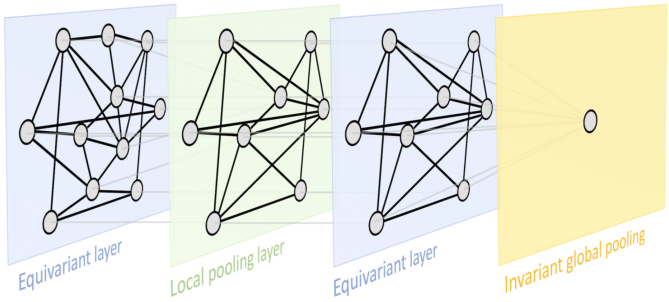

Fig. 1. Generic structure of Equivariant Neural Networks (from [6]): ENN typically share similar topology with conventional CNN, with (generalized) equivariant convolution layers and (local or global) pooling mechanisms.

2.1 Symmetries Representation

In the following, symmetries will be represented by the invariance/equivariance of a function to the action of a Lie group on the input data. More precisely, a group G is acting on a set S if there exists a map $\circ : \ G \times S \to S$ which is compatible with the group law $\circ$ in the sense that $h \circ (g \circ S) = (hg) \circ S, \forall g, h \in G$. For two sets X and Y on which a group G acts respectively with $\circ_X$ and $\circ_Y$, a function $f : \ X \to Y$ is said to be G-equivariant (or G-covariant) if $\forall x \in X$ and $\forall g \in G$, $f\left(g \circ_X x\right) = g \circ_Y f\left(x\right)$. Similarly, $f : \ X \to Y$ is said to be G-invariant if $f\left(g \circ_X x\right) = f\left(x\right)$, $\forall x \in X$ and $\forall g \in G$. Hence, G-invariance is a special case of $G-$equivariance, for which the group action $\circ_Y$ is trivial. To illustrate the above, we consider the set I_2 of $2-$dimensional gray scale images that we represent by continuous functions $f : \ \mathbb{R}^2 \to [-1, 1]$, where $f\left(x, y\right)$ represents the value of the renormalized pixel at position (x, y). Examples of Lie groups acting on the set I_2 include the translation group $\mathbb{R}^2$, the rotation group SO(2) and the special euclidean group SE(2). An image classifier is typically expected to be invariant with respect to such action, while segmentation algorithms should be equivariant, as illustrated on Figure 2.

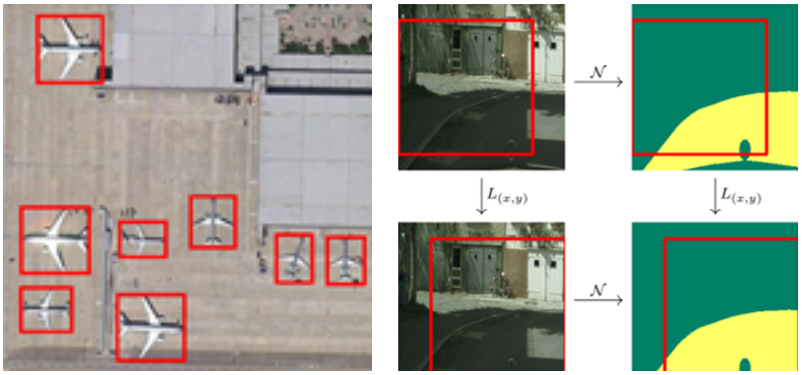

Fig. 2. Left (image from [27]): the planes should be classified as such regardless of their orientation. Corresponding algorithms are in particular expected to be invariant to the rotation group with this respect. Right (image from [15]): commutative diagram representing the expected equivariant behavior of segmentation algorithms.

2.2 From CNN to G-CNN

CNN [19] are by-design well adapted to computer vision tasks as the underlying 2d convolution operators are equivariant to translation, allowing for efficient weights sharing for features extraction. Unfortunately, the equivariance property of conventional CNN is limited to translation, which is a caveat even for planar images for which equivariance to rotation and scaling would be a beneficial for the inner representation of the data.

Group-Convolutional Neural Networks (G-CNN) have initially been introduced in 2016 by the seminal work [8] as a generalization of CNN by introducing group-based convolution operators (see Section 2.3) to achieve equivariance with respect to the action of finite groups and have further been generalized to more generic actions. In particular, [10] proposed a sound theory for the case of the transitive action of a compact group on its homogeneous space by leveraging on group representation theory. More recently, [13] introduced a very generic approach providing equivariance to the action of any Lie group with a surjective exponential map and which is applicable to any input data representable by a function defined on a smooth manifold and valued in a vectorial space.

2.3 Group-Based Convolution

A natural way to achieve equivariance with respect to the action of a given group G is to generalize the 2d planar convolution. More precisely, we consider that the kernel κ and feature map f are signals on the group G valued in a vectorial space V such that $\kappa, f : G \to V$ and we define the generalized convolution operator on the group G by writing, $\forall g \in G$:

$$(\kappa \star f)(g) = \int_G \kappa\left(h^{-1}g\right) f(h) d\mu^G(h) \tag{1}$$

where μ^G is the Haar measure of the group G. Denoting L_h the left shift operator such that $\forall h \in G$, $\forall g \in G$ and $\forall \phi : G \to V$, $L_h\phi(g) = \phi\left(h^{-1}g\right)$, the operator defined by equation (1) is equivariant with respect tot the action of G in the sense that $L_h(\kappa \star f)(g) = (\kappa \star L_h f)(g)$. Specializing with $G = \mathbb{R}^2$ allows retrieving the usual 2d convolution operator. Group-based convolution operators can then be used within G-CNN architectures to instanciate equivariant layers of the ENN blueprint shown on Figure 1.

2.4 Adversarial Robustness

In addition to providing by-design geometrical robustness, preliminary studies have shown that G-CNN also appear more robust to adversarial attacks than conventional architectures in the context of image classification. In [11], SO(2) is considered and we have extended the study to the SE(2) case by leveraging on the corresponding ENN implementation of [16]. Fig 3 in particular shows that, for a given number of iterations, the adversarial perturbations corresponding to DeepFool adversarial MNIST samples [24] for a SE(2)-ENN are higher in expectation than those corresponding to usual approaches (MLP and CNN).

2.5 Applications

ENN have already found a wide range of applications, including in Computer Vision [16,30], for Graph and Point Cloud processing [14,26], Simulation and Trajectory prediction [13,29], in Reinforcement learning [25] and for Time Series

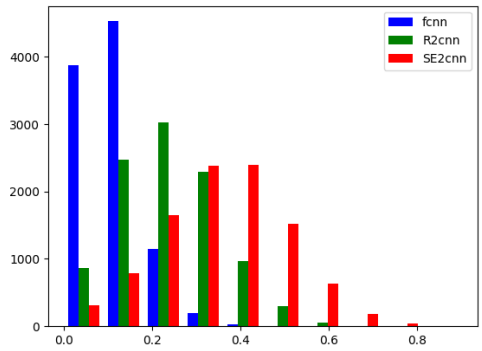

Fig. 3. Distribution of the size (ℓ_2-norm) of the adversarial perturbations obtained with the DeepFool attack for a fixed number of iterations, for a SE(2)-ENN (SE2cnn), a conventional CNN (R2cnn) and a simple MLP (fcnn) with roughly the same number of parameters and trained on the MNIST dataset.

Analysis [18,17], all of these areas being of interest in the context of building safe and secure AI-based Defense systems. More precisely, we highlight below examples of possible Defense applications:

- **Robust remote sensing image processing** such as the detection of planes from satellite images by exploiting roto-translation and scale equivariance [27].
- **Native omnidirectional image processing** by using 3d rotation equivariance as to benefit from the native spherical geometry of the input data (e.g., obtained with bidirectional fish-eye lenses) and to avoid dealing with the distortion effects induced by planar projection methods [9].
- **Efficient simulation** of Partial Differential Equations (PDE) such as Maxwell and Navier Stokes equations by building ENN equivariant to the symmetry group of the underlying PDE [29], allowing for faster convergence and better stability than finite difference/element methods.
- **Robust radar Doppler signal processing** by representing the signal as complex covariance matrices and combining equivariance with hyperbolic embedding to increase the robustness to real-world perturbations such as thermal noise [18,17].

3 Lie Group Statistics and Machine Learning

Lie Group Statistics and Machine Learning is a very promising aera of research with well identified Defense applications such as robotics, radar Doppler signal processing and kinematics recognition [4], and which anchors in deep theoretical results [5,2,3,21,22]. Also, the practical implementation of these approaches

becomes more and more convenient with the emergence of dedicated python packages such as Geomstats [23].

More precisely, the geometric models of information theory, statistical physics and machine learning inference share common structures as was illustrated at the Ecole de Physique des Houches in July 2020, SPIGL'20 [1]. Information Geometry has introduced natural learning gradients that are invariant to information coding via the Fisher metric. These schemes have been extended to differential manifolds or Lie groups when symmetries exist by extending the notion of Fisher metric to Koszul-Souriau metric and the definition of Entropy as invariant Casimir function in coadjoint representation. To define statistics on Lie Group, we have to consider the associated Symplectic homogeneous manifold given by the Kirillov-Kostant-Souriau 2-form, associated to the Lie group coadjoint orbits, and to define extension of Gauss density as Gibbs density. This model has been introduced by Jean-Marie Souriau and named "Lie groups thermodynamics". In the Souriau model, Fisher-Koszul metric is given by:

$$I(\beta) = -\frac{\partial^2 \Phi}{\partial \beta^2} \text{ with } \Phi(\beta) = -\log \int\limits_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega \text{ and } U : M \to \mathfrak{g}^* \tag{2}$$

where the Entropy is given by Legendre transform: $S(Q) = \langle Q, \beta \rangle - \Phi(\beta)$ with $Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in \mathfrak{g}^*$ and $\beta = \frac{\partial S(Q)}{\partial Q} \in \mathfrak{g}$, where β is a "geometric" (Planck) temperature, element of Lie algebra $\mathfrak{g}$ of the group, and Q is a "geometric" heat, element of the dual space of the Lie algebra $\mathfrak{g}^*$ of the group. Souriau defined a Gibbs density that is covariant under the action of the group:

$$p_{Gibbs}(\xi) = e^{\Phi(\beta) - \langle U(\xi), \beta \rangle} = \frac{e^{-\langle U(\xi), \beta \rangle}}{\int\limits_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega} \tag{3}$$

with $\Phi(\beta) = -\log \int\limits_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega$. We can express the Gibbs density with respect to Q by inverting the relation $Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta)$. Then $p_{Gibbs,Q}(\xi) = e^{\Phi(\beta) - \left\langle U(\xi), \Theta^{-1}(Q) \right\rangle}$ with $\beta = \Theta^{-1}(Q)$.

We can then beneficiate of different tools based on Souriau Lie Groups Thermodynamics and Kirillov Representation Theory for:

- **Supervised Machine Learning**: Maximum Entropy covariant density on co-adjoint orbits through moment map and extension of natural gradient on Lie algebra for Lie groups Machine Learning.
- **Non-Supervised Machine Learning**: Extension of Mean/Median on Lie group by Fréchet definition of geodesic barycenter on Souriau-Fisher Metric Space, solved by Karcher Flow, and extension of Mean-Shift for homogeneous symplectic manifold and Souriau-Fisher Metric Space.

4 Conclusion

We have given some general background about ENN and explained their rational in the context of achieving robustness with respect to geometrical transforms. We

have also shed light on Lie Group Statistics and Machine Learning techniques and their interest for exploiting the native geometry of the input data. Both technologies have a wide range of applications for the Defense industry and we generally believe that exploiting data geometry and the underlying symmetries is key to the design of efficient, reliable and robust AI-based Defense systems.

References

1. Barbaresco, F., Nielsen, F.: Geometric Structures of Statistical Physics, Information Geometry, and Learning: SPIGL'20, Les Houches, France, July 27–31. Springer Proceedings in Mathematics & Statistics, Springer International Publishing (2021), https://books.google.fr/books?id=jbFqzgEACAAJ
2. Barbaresco, F.: Archetypal model of entropy by poisson cohomology as invariant casimir function in coadjoint representation and geometric fourier heat equation. In: Nielsen, F., Barbaresco, F. (eds.) Geometric Science of Information. pp. 417–429. Springer International Publishing, Cham (2021)
3. Barbaresco, F.: Gaussian distributions on the space of symmetric positive definite matrices from souriau's gibbs state for siegel domains by coadjoint orbit and moment map. In: Nielsen, F., Barbaresco, F. (eds.) Geometric Science of Information. pp. 245–255. Springer International Publishing, Cham (2021)
4. Barbaresco, F.: Lie group statistics and lie group machine learning based on souriau lie groups thermodynamics & koszul-souriau-fisher metric: New entropy definition as generalized casimir invariant function in coadjoint representation. Entropy **22**(6) (2020). https://doi.org/10.3390/e22060642, https://www.mdpi.com/1099-4300/22/6/642
5. Barbaresco, F., Gay-Balmaz, F.: Lie group cohomology and (multi)symplectic integrators: New geometric tools for lie group machine learning based on souriau geometric statistical mechanics. Entropy **22**(5) (2020). https://doi.org/10.3390/e22050498, https://www.mdpi.com/1099-4300/22/5/498
6. Bronstein, M.M., Bruna, J., Cohen, T., Veličković, P.: Geometric deep learning: Grids, groups, graphs, geodesics, and gauges (2021)
7. Chen, S., Dobriban, E., Lee, J.H.: A group-theoretic framework for data augmentation. Journal of Machine Learning Research **21**(245), 1–71 (2020), http://jmlr.org/papers/v21/20-163.html
8. Cohen, T., Welling, M.: Group equivariant convolutional networks. In: Balcan, M.F., Weinberger, K.Q. (eds.) Proceedings of The 33rd International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 48, pp. 2990–2999. PMLR, New York, New York, USA (20–22 Jun 2016), http://proceedings.mlr.press/v48/cohenc16.html
9. Cohen, T.S., Geiger, M., Koehler, J., Welling, M.: Spherical cnns (2018)
10. Cohen, T.S., Geiger, M., Weiler, M.: A general theory of equivariant cnns on homogeneous spaces. In: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems. vol. 32, pp. 9145–9156. Curran Associates, Inc. (2019), https://proceedings.neurips.cc/paper/2019/file/b9cfe8b6042cf759dc4c0cccb27a6737-Paper.pdf
11. Dumont, B., Maggio, S., Montalvo, P.: Robustness of rotation-equivariant networks to adversarial perturbations (2018)

12. Elesedy, B., Zaidi, S.: Provably strict generalisation benefit for equivariant models (2021)
13. Finzi, M., Stanton, S., Izmailov, P., Wilson, A.G.: Generalizing convolutional neural networks for equivariance to lie groups on arbitrary continuous data (2020)
14. Fuchs, F.B., Worrall, D.E., Fischer, V., Welling, M.: Se(3)-transformers: 3d roto-translation equivariant attention networks (2020)
15. Gerken, J.E., Aronsson, J., Carlsson, O., Linander, H., Ohlsson, F., Petersson, C., Persson, D.: Geometric deep learning and equivariant neural networks (2021)
16. Lafarge, M.W., Bekkers, E.J., Pluim, J.P.W., Duits, R., Veta, M.: Roto-translation equivariant convolutional networks: Application to histopathology image analysis (2020)
17. Lagrave, P.Y., Cabanes, Y., Barbaresco, F.: "su(1,1) equivariant neural networks and application to robust toeplitz hermitian positive definite matrix classification". In: Nielsen, F., Barbaresco, F. (eds.) Geometric Science of Information. pp. 577–584. Springer International Publishing, Cham (2021)
18. Lagrave, P.Y., Cabanes, Y., Barbaresco, F.: An equivariant neural network with hyperbolic embedding for robust doppler signal classification. In: 2021 21st International Radar Symposium (IRS). pp. 1–9 (2021). https://doi.org/10.23919/IRS51887.2021.9466226
19. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998). https://doi.org/10.1109/5.726791
20. Li, F., Zhang, L., Zhang, Z.: Lie Group Machine Learning. De Gruyter (2018). https://doi.org/doi:10.1515/9783110499506, `https://doi.org/10.1515/9783110499506`
21. Marle, C.M.: On Gibbs States of Mechanical Systems with Symmetries. Journal of Geometry and Symmetry in Physics **57**(none), 45 – 85 (2020). https://doi.org/10.7546/jgsp-57-2020-45-85, `https://doi.org/10.7546/jgsp-57-2020-45-85`
22. Marle, C.M.: On gibbs states of mechanical systems with symmetries (2021)
23. Miolane, N., et al.: Geomstats: A python package for riemannian geometry in machine learning. Journal of Machine Learning Research **21**(223), 1–9 (2020), `http://jmlr.org/papers/v21/19-027.html`
24. Moosavi-Dezfooli, S.M., Fawzi, A., Frossard, P.: Deepfool: A simple and accurate method to fool deep neural networks. pp. 2574–2582 (06 2016). https://doi.org/10.1109/CVPR.2016.282
25. van der Pol, E., Worrall, D.E., van Hoof, H., Oliehoek, F.A., Welling, M.: Mdp homomorphic networks: Group symmetries in reinforcement learning (2021)
26. Satorras, V.G., Hoogeboom, E., Welling, M.: E(n) equivariant graph neural networks (2021)
27. Shamsolmoali, P., Zareapoor, M., Chanussot, J., Zhou, H., Yang, J.: Rotation equivariant feature image pyramid network for object detection in optical remote sensing imagery (2021)
28. Shorten, C., Khoshgoftaar, T.: A survey on image data augmentation for deep learning. Journal of Big Data **6** (07 2019). https://doi.org/10.1186/s40537-019-0197-0
29. Wang, R., Walters, R., Yu, R.: Incorporating symmetry into deep dynamics models for improved generalization (2021)
30. Weiler, M., Forré, P., Verlinde, E., Welling, M.: Coordinate independent convolutional networks – isometry and gauge equivariant convolutions on riemannian manifolds (2021)