HIGHER ORDER GAUGE EQUIVARIANT CONVOLUTIONS FOR NEURODEGENERATIVE DISORDER CLASSIFICATION

Gianfranco Cortés[†] *Yue Yu* *Robin Chen*
Melissa Armstrong *David Vaillancourt* *Baba C. Vemuri*

University of Florida

ABSTRACT

Diffusion MRI (dMRI) has shown significant promise in capturing subtle changes in neural microstructure caused by neurodegenerative disorders. In this paper, we propose a novel end-to-end compound architecture for processing raw dMRI data. It consists of a 3D convolutional kernel network (CKN) that extracts macro-architectural features across voxels and a gauge equivariant Volterra network (GEVNet) on the sphere that extracts micro-architectural features from within voxels. The use of higher order convolutions enables our architecture to model spatially extended nonlinear interactions across the applied diffusion-sensitizing magnetic field gradients. The compound network is globally equivariant to 3D translations and locally equivariant to 3D rotations. We demonstrate the efficacy of our model on the classification of neurodegenerative disorders.

***Index Terms*—** diffusion MRI, gauge equivariance, neurodegenerative diseases, higher order convolutions

1. INTRODUCTION

Alzheimer's disease (AD), Dementia with Lewy bodies (DLB), and Parkinson's disease (PD) are the most common causes of neurodegenerative dementia, but effective diagnosis and differentiation between the three remains a major challenge. As many as two out of every three cases of DLB are either missed entirely or misdiagnosed, most commonly as AD [1]. To the best of our knowledge, no end-to-end deep networks have been reported in the literature specifically for discriminating between AD, DLB, and PD.

Diffusion MRI is a non-invasive imaging modality that can capture variations in neural microstructure. Recent deep learning works have demonstrated the utility of dMRI in the analysis of Parkinson's disease, e.g. [2]. However, most of these methods do not apply directly to raw dMRI data but instead apply to derived representations such as diffusion tensors and fiber orientation distribution functions.

A raw dMRI signal (single-shell acquisition) is modeled as a function $f : \mathbf{R}^3 \times S^2 \to \mathbf{R}$, i.e. there exists a 3D volume $f_g : \mathbf{R}^3 \to \mathbf{R}$ for every diffusion-sensitizing magnetic gradient direction (or b-vector) $g \in S^2$. While traditional convolutional neural networks (CNNs) can be artificially adapted to exploit raw dMRI data, doing so requires ignoring the underlying geometry and thus loses potentially valuable information. A more effective approach is to embrace the tenets of geometric deep learning (GDL), a subfield that generalizes convolutions to non-Euclidean settings such as the sphere S^2.

Several GDL-inspired works acting on raw diffusion data exist in the literature [3, 4, 5]. In this work, however, we are interested in extending the success of higher order convolutions in dMRI witnessed in [2, 6]. We do this by developing a higher order analogue of the gauge equivariant (GE) convolution introduced in [7], an extremely general convolution operation that is valid on arbitrary Riemannian manifolds. Despite their theoretical generality, GE convolutions exhibit practical benefits over Fourier-theoretic group convolutions, as they decouple the spatial resolution from the feature map bandwidth.

The key contributions of this work are: (1) a higher order generalization of the GE convolution, (2) an implementation on S^2, dubbed the GEVNet, and (3) experimental verification of the utility of a novel compound architecture, CKN2 + GEVNet, for classifying neurodegenerative diseases.

2. BACKGROUND

2.1. First Order Gauge Equivariant Convolutions

We now briefly review the theory of GE convolutions and refer the reader to [7, 8] for a more detailed treatment.

Let M be a 2-dimensional Riemannian manifold with structure group $\mathsf{SO}(2)$, the group of 2D rotations. In the GE setting, feature maps become tensor fields on M (e.g. vector fields). A feature map is coordinatized w.r.t. a gauge, a local choice of coordinate frame. A GE convolution maps an input feature map f_{in} of type ρ_{in} to an output feature map f_{out} of type ρ_{out}, where ρ_{in} and ρ_{out} are group representations that encode the transformation behavior of the tensor components under a change of gauge. Let f_{in} be a feature map of type ρ_{in} and $K : \mathbf{R}^2 \to \mathbf{R}^{d_{\text{out}} \times d_{\text{in}}}$ a matrix-valued filter where d_{in}

[*]This research was in part supported by NIH NIA and NINDS RO1 NS121099 to Vemuri and the McKnight Doctoral Fellowship to Cortés.

[†]Correspondence to: gcortes@ufl.edu

and d_{out} are the dimensions of the input and output tensors, respectively. Letting $q_v := \exp_p w_p v$ (exp denotes the Riemannian exponential map and w_p a gauge), the GE convolved feature map $f_{\text{out}} = K \star f_{\text{in}}$ is given pointwise w.r.t. w_p by

$$f_{\text{out}}(p) := \int_{\mathbf{R}^2} K(v)\rho_{\text{in}}(t_{p\leftarrow q_v}) f_{\text{in}}(q_v)\, dv, \tag{1}$$

where $t_{p\leftarrow q_v}$ denotes the $\mathsf{SO}(2)$-valued gauge transformation taking the frame on q_v (after parallel transport to p) to the frame on p. Eqn. 1 is equivariant to a change of gauge at p if and only if K is $\mathsf{SO}(2)$-*steerable*, i.e. K satisfies

$$K(t^{-1}v) = \rho_{\text{out}}(t^{-1})K(v)\rho_{\text{in}}(t) \tag{2}$$

for all $t \in \mathsf{SO}(2)$ and $v \in \mathbf{R}^2$. A critical result of [8] shows that, in the case where $M = S^2$, GE convolutions with $\mathsf{SO}(2)$-steerable filters are equivariant to rotations $\phi \in \mathsf{SO}(3)$.

2.2. The Volterra Series

The traditional convolution operator admits a natural generalization called the *Volterra series* (or *expansion*) [9], which in practice is truncated at some specified order. For example, given a 1D signal $f : \mathbf{R} \to \mathbf{R}$, its second order Volterra expansion is given by $V^2[f] = h^{(0)} + h^{(1)} \star f + h^{(2)} \star f$, where $h^{(0)}$ is a constant, $h^{(1)} \star f$ is the usual (first order) convolution, and

$$(h^{(2)} \star f)(x) = \iint h^{(2)}(\tau_1, \tau_2) f(x - \tau_1) f(x - \tau_2)\, d\tau_1 d\tau_2. \tag{3}$$

Each $h^{(k)}$ is a learnable filter taking k arguments. The intuition behind Eqn. 3 is that the second order term is aggregating all pairwise interactions between neighbors within the filter support. In general, the higher order terms for $k > 1$ can model *spatially extended* nonlinear interactions across a receptive field, a task that *pointwise* nonlinearities such as the ReLU are less suited for.

3. METHODOLOGY

3.1. Second Order Gauge Equivariant Convolutions

Here we propose a higher order analogue of the first order GE convolution seen in Eqn. 1, inspired by the Volterra expansion defined in Eqn. 3. We will only be concerned with second order expansions for the remainder of this work. *We omit all proofs due to space constraints*. Let M be as defined in section 2.1.

Definition 3.1. A second order filter $K^{(2)}$ of type $(\rho_{\text{out}}, \rho_{\text{in}})$ is a smooth map

$$K^{(2)} : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^{d_{\text{out}} \times d_{\text{in}}^2}. \tag{4}$$

Definition 3.2. A second order filter $K^{(2)}$ of type $(\rho_{\text{out}}, \rho_{\text{in}})$ is said to be $\mathsf{SO}(2)$-steerable if and only if it satisfies a second order steerability constraint given by

$$K^{(2)}(t^{-1}v_1, t^{-1}v_2) = \rho_{\text{out}}(t^{-1})K^{(2)}(v_1, v_2)\rho_{\text{in}}^{\otimes 2}(t) \tag{5}$$

for all $t \in \mathsf{SO}(2)$ and $v_1, v_2 \in \mathbf{R}^2$. Here, $\rho_{\text{in}}^{\otimes 2}$ is the tensor product representation given by $\rho_{\text{in}}^{\otimes 2}(t) = \rho_{\text{in}}(t) \otimes \rho_{\text{in}}(t)$.

Definition 3.3. Let $K^{(1)}$ be a first order $\mathsf{SO}(2)$-steerable filter (see Eqn. 2) and let $K^{(2)}$ be a second order $\mathsf{SO}(2)$-steerable filter, both of type $(\rho_{\text{out}}, \rho_{\text{in}})$. Letting $q_v := \exp_p w_p v$, the second order GE expansion of a feature map f_{in} is given by $V_{\text{GE}}^2[f_{\text{in}}] := K^{(1)} \star f_{\text{in}} + K^{(2)} \star f_{\text{in}}$, where $K^{(1)} \star f_{\text{in}}$ is the first order GE convolution in Eqn. 1 and

$$(K^{(2)} \star f_{\text{in}})(p) := \int_{\mathbf{R}^2}\int_{\mathbf{R}^2} K^{(2)}(v_1, v_2) \\ \rho_{\text{in}}(t_{p\leftarrow q_{v_1}}) f_{\text{in}}(q_{v_1}) \otimes \rho_{\text{in}}(t_{p\leftarrow q_{v_2}}) f_{\text{in}}(q_{v_2})\, dv_1 dv_2. \tag{6}$$

Note how Eqn. 6 reduces to the second order Volterra expansion in Eqn. 3 in the case of scalar-valued feature maps on a Euclidean space (save the zeroth order bias term $h^{(0)}$). We now present a theorem on the equivariance of the operator V_{GE}^2. Such a result is necessary because (a) we would like for $V_{\text{GE}}^2[f_{\text{in}}]$ to be independent of the coordinatization of f_{in} and (b) we want V_{GE}^2 to further enjoy $\mathsf{SO}(3)$-equivariance in the case where $M = S^2$ (see the last sentence of section 2.1).

Theorem 3.4. *V_{GE}^2 is gauge equivariant. That is, if $f_{\text{in}}(p)$ is coordinatized with respect to $\widetilde{w}_p = w_p \circ t$, where $t \in \mathsf{SO}(2)$, then $V_{\text{GE}}^2[f_{\text{in}}](p)$ transforms as $\rho_{\text{out}}(t^{-1})V_{\text{GE}}^2[f_{\text{in}}](p)$ when coordinatized with respect to w_p.*

The main difficulty of applying Eqn. 6 in practice is constructing filters that satisfy the steerability constraints in Eqns. 2 and 5. However, since our implementation in section 3.2 will only rely on zeroth and first order tensors (i.e. scalars and tangent vectors), we can make do with the following sufficient condition for second order $\mathsf{SO}(2)$-steerability.

Proposition 3.5. *Let $K_{01}^{(1)}$ be a first order $\mathsf{SO}(2)$-steerable filter of type (ρ_0, ρ_1) and $K_{11}^{(1)}$ a first order $\mathsf{SO}(2)$-steerable filter of type (ρ_1, ρ_1), where ρ_0 is the trivial representation (i.e. $\rho_0(t) = 1$) and ρ_1 is the standard representation mapping $t \in \mathsf{SO}(2)$ to the usual 2×2 rotation matrix. Then, a second order filter of the form*

$$K_{11}^{(2)}(v_1, v_2) := K_{11}^{(1)}(v_1) \otimes K_{01}^{(1)}(v_2) \tag{7}$$

is $\mathsf{SO}(2)$-steerable (i.e. satisfies Eqn. 5) of type (ρ_1, ρ_1).

The previous proposition enables us to construct second order $\mathsf{SO}(2)$-steerable basis filters using the first order $\mathsf{SO}(2)$-steerable solutions to Eqn. 2 derived in [10]. Note that the proposition can be modified to incorporate second order $\mathsf{SO}(2)$-steerable filters of types (ρ_0, ρ_1), (ρ_1, ρ_0), and (ρ_0, ρ_0). We then learn a linear combination of the generated second order basis filters.

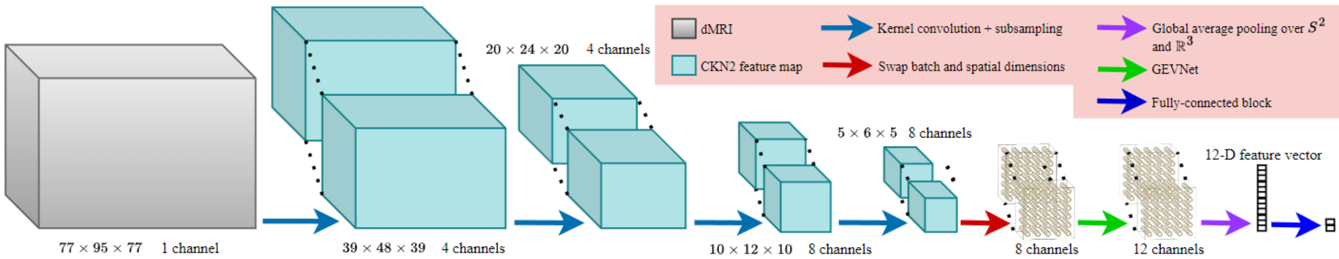

Fig. 1. A schematic of the CKN2 + GEVNet model. Letting $\mathbf{GEVConv}(c_{\text{in}}^{\rho_{\text{in}}}, c_{\text{out}}^{\rho_{\text{out}}})$ denote a GEVNet convolution layer taking in c_{in} feature maps of type ρ_{in} and outputting c_{out} feature maps of type ρ_{out}, our implemented GEVNet is of the form $\mathbf{GEVConv}(8^{\rho_0}, 8^{\rho_0 \oplus \rho_1}) \to \mathbf{GEVConv}(8^{\rho_0 \oplus \rho_1}, 12^{\rho_0 \oplus \rho_1}) \to \mathbf{GEVConv}(12^{\rho_0 \oplus \rho_1}, 12^{\rho_0})$.

3.2. The GEVNet

The GEVNet is an instantiation of Eqn. 6 on $M = S^2$. It computes second order GE convolutions with hidden feature maps of type $\rho_0 \oplus \rho_1$, where ρ_0 and ρ_1 are as in Prop. 3.5. Due to memory constraints, we assume that pairwise tensor products only occur across features of the same order. Explicitly, if $f(q_1) = s_1 \oplus r_1$ and $f(q_2) = s_2 \oplus r_2$, where s_i is a scalar and r_i is a tangent vector, then we stipulate that $\boxed{f(q_1) \otimes f(q_2) := (s_1 \otimes s_2) \oplus (r_1 \otimes r_2)}$, i.e. we assume vanishing cross terms. The second order GE convolutions are interleaved with regular nonlinearities as described in [11].

4. EXPERIMENTS

4.1. Dataset and Preprocessing

Our dMRI data pool consisted of a cohort of 112 AD, 85 DLB, and 436 PD patient brain scans. The scans were eddy current and motion corrected using FSL [12]. Since the data were pooled from different MR scanners possessing distinct acquisition parameters (see URLs provided in section 7), the scans underwent a retrospective harmonization using [13]. Next, the scans were affinely registered to a common MNI template and downsampled to a voxel size of 2 mm^3. Finally, the image intensities at each voxel were interpolated and resampled (constrained by antipodal symmetry of the diffusion signal) onto a Healpix grid with 192 grid points. For each classification problem (see 4.3), we mitigate class imbalance in the training set by augmenting the deficient class using *mixup* [14] after uniformly sampling a 20% test set.

4.2. Network Architecture

Our proposed architecture for processing raw dMRI scans consists of a macro-architectural module that extracts features across voxels, followed by a micro-architectural module that extracts features from within voxels. The macro-architectural module is a convolutional kernel network (CKN) [15] on $\mathbf{R}^3$. The details of the CKN are outside the scope of our present work, but it suffices to understand it as an approximation to a standard CNN that can be easily modified to efficiently compute higher order convolutions on $\mathbf{R}^3$. This is done by modifying the degree of a polynomial RKHS kernel, e.g. a linear kernel approximates first order convolutions while a quadratic kernel approximates second order convolutions. The micro-architectural module is a spherical CNN.

A single dMRI scan is an array of shape $192 \times 1 \times 77 \times 95 \times 77$, corresponding to 192 b-vectors, one input channel, and three spatial dimensions. Recall there is a 3D volume for each b-vector. Since the b-vectors populate the batch dimension, the CKN computes inter-voxel features for each 3D volume in isolation. The resulting feature map has shape $192 \times 8 \times 5 \times 6 \times 5$. At this point, we swap the spatial dimensions with the batch dimension, since now the b-vectors become precisely the grid positions on S^2. The spherical CNN's output is pooled over $\mathbf{R}^3$ and S^2 to produce a final feature vector. See Fig. 1 for a schematic. Although this two-part architecture is not invariant to the joint action of $\mathsf{SE}(3)$, there is evidence [5] to suggest that this is an unnecessary requirement since all scans are registered prior to downstream processing.

4.3. Ablation Study

We test the effect of higher order convolutions on three neurodegenerative disease classification problems (Table 1). We ablate on the order of the inter-voxel convolution (CKN) by varying the degree of a polynomial kernel, and on the order of the intra-voxel convolution (spherical CNN) by exchanging a second order GEVNet with a first order DeepSphere (DS) [16]. The notation CKN1 and CKN2 refer to a CKN with linear and quadratic kernels, respectively. We also include two baselines: a lone CKN2 that ignores b-vector geometry, and a lone GEVNet that ignores diffusion across voxels.

We control for parameter counts to eliminate their possibility as a confound. Each CKN has a *fixed* architecture consisting of 3164 learnable parameters. Both the GEVNet and the DS consist of three convolution layers, with the GEVNet having 12424 parameters and the DS having 13696 parameters. All other optimization-related hyperparameters are kept the same across classification tasks. For a fixed choice of

Table 1. Test accuracies (%) on three neuroimaging classification problems, averaged over five runs.

Architecture	AD v. DLB	AD v. PD	DLB v. PD
CKN2	86.76	90.32	98.06
GEVNet	77.08	72.36	81.12
CKN1 + DS	86.76	79.16	89.24
CKN1 + GEVNet	90.34	86.30	93.30
CKN2 + DS	89.82	95.07	**98.46**
CKN2 + GEVNet	**92.86**	**98.36**	98.27

Table 2. Extra performance metrics for the CKN2 + GEVNet.

Metric (%)	AD v. DLB	AD v. PD	DLB v. PD
Sensitivity	91.30	98.18	98.82
Specificity	95.00	98.41	98.16

CKN, we find that an accompanying GEVNet outperforms (or is highly comparable to) its DS counterpart by a significant margin. This is noteworthy given that the GEVNet is at an over 1k parameter disadvantage against the DS. Conversely, for a fixed choice of spherical CNN, we find that replacing the accompanying CKN's linear kernel with a quadratic kernel yields a boost in performance. We see that CKN2 + GEVNet is the overall best classifier, as it performs second order convolutions both across and within voxels. Additional supporting metrics for this classifier are provided in Table 2.

5. CONCLUSION

In this paper, we generalized the first order GE convolution to its higher order analogue. Our resulting implementation on S^2, the GEVNet, was applied as part of a compound architecture for the classification of neurodegenerative disorders. Our experiments demonstrate the importance of considering higher order convolutions in settings exhibiting spatially extended nonlinear interactions (e.g. diffusion of water molecules in section 4.3) which are difficult to capture by solely relying on pointwise nonlinearities.

6. REFERENCES

[1] A.J. Thomas, J.P. Taylor, et al., "Development of assessment toolkits for improving the diagnosis of the LBDs: feasibility study within the DIAMOND Lewy study," *Intl. J. of Geriatric Psych.*, vol. 32(12), pp. 1280–1304, 2017.

[2] J. Bouza, C. Yang, et al., "A higher order manifold-valued convolutional neural network with applications to dMRI processing," in *IPMI 2021*, pp. 304–317.

[3] S. Sedlar, A. Alimi, et al., "A spherical convolutional neural network for white matter structure imaging via dMRI," in *MICCAI*, 2021, pp. 529–539.

[4] P. Müller, V. Golkov, et al., "Rotation-equivariant deep learning for diffusion MRI," *arXiv:2102.06942*, 2021.

[5] R. Liu, F. Lauze, et al., "SE(3) Group Conv. NN and a Study on Group Convs. and Equivariance for DWI Segmentation," 2023, Preprint.

[6] M. Banerjee, R. Chakraborty, et al., "VolterraNet: A higher order convolutional network with group equivariance for homogeneous manifolds," *IEEE TPAMI*, vol. 44, no. 2, pp. 823–833, 2020.

[7] T.S. Cohen, M. Weiler, et al., "Gauge equivariant convolutional networks and the icosahedral CNN," in *ICML*. PMLR, 2019, pp. 1321–1330.

[8] M. Weiler, P. Forré, et al., "Coordinate Independent Convolutional Networks–Isometry and Gauge Equivariant Convolutions on Riemannian Manifolds," *arXiv:2106.06020*, 2021.

[9] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations*, 1930.

[10] M. Weiler and G. Cesa, "General E(2)-equivariant steerable CNNs," *NIPS*, vol. 32, 2019.

[11] P. de Haan, M. Weiler, et al., "Gauge Equivariant Mesh CNNs: Anisotropic convolutions on geometric graphs," in *ICLR*, 2021.

[12] J.L.R. Andersson and S.N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging," *Neuroimage*, vol. 125, pp. 1063–1078, 2016.

[13] S.C. Karayumak, S. Bouix, et al., "Retrospective harmonization of multi-site diffusion MRI data acquired with different acquisition parameters," *Neuroimage*, vol. 184, pp. 180–200, 2019.

[14] H. Zhang, M. Cisse, et al., "mixup: Beyond Empirical Risk Minimization," in *ICLR*, 2018.

[15] J. Mairal, "End-to-end kernel learning with supervised convolutional kernel networks," *NIPS*, vol. 29, 2016.

[16] M. Defferrard, M. Milani, et al., "DeepSphere: a graph-based spherical CNN," in *ICLR*, 2020.

7. COMPLIANCE WITH ETHICAL STANDARDS

The dMRI data used in this manuscript was acquired retrospectively and made openly available for access via (1) `https://pdbp.ninds.nih.gov` (PDBP), (2) `https://1floridaadrc.org` (ADRC), and (3) `https://www.ppmi-info.org` (PPMI). Ethical approval was not required as confirmed by the license attached with the open access data.