GM-GAN: Geometric Generative Models based on Morphological Equivariant PDEs and GANs

Hadji S El, Thierno Fall, Alioune Mbengue, Mohamed Daoudi

GM-GAN: Geometric Generative Models based on Morphological Equivariant PDEs and GANs

El Hadji S. Diop[1], Thierno Fall[1], Alioune Mbengue[1,2], and Mohamed Daoudi[3]

[1] NAGIP-Nonlinear Analysis and Geometric Information Processing Group, Department of Mathematics, University Iba Der Thiam, Thies BP 967, Senegal
ehsdiop@hotmail.com,thiernofall571@gmail.com

[2] Department of Mathematics and Computer Science, University Cheikh Anta Diop, Dakar, Senegal
99aliou@gmail.com

[3] IMT Nord Europe, Univ. Lille, Centre for Digital Systems, F-59000 Lille, France
Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France
mohamed.daoudi@imt-nord-europe.fr

Abstract. This work deals with image generation, two main problems are addressed: (*i*) improvements of specific feature extraction while accounting at multiscale levels intrinsic geometric features, and (*ii*) equivariance of the network for reducing the complexity and providing a geometric interpretability. We propose a geometric generative model based on an equivariant partial differential equation (PDE) for group convolution neural networks (G-CNNs), so called PDE-G-CNNs, built on morphology operators and generative adversarial networks (GANs). The proposed geometric morphological GAN model, termed as GM-GAN, is obtained thanks to morphological equivariant convolutions in PDE-G-CNNs. GM-GAN is evaluated qualitatively and quantitatively using FID on MNIST and RotoMNIST, preliminary results show noticeable improvements compared classical GAN.

Keywords: PDEs, Equivariance, Morphological operators, Riemannian manifolds, Lie group, Symmetries, CNNs.

1 Introduction

Significant advances in deep learning progress are attributed to CNNs [23]. Despite its successful applications in many real life problems, it is still not very clear why deep learning techniques work. Pursuing this goal, many works attempt to give an answer to this so challenging question by setting mathematical frameworks that underlie the process. A promising direction is to consider symmetries as a fundamental design principle for network architectures. Among noticeable properties in CNNs, the equivariance concerning translations played an important role. Equivariance means that the operation of performing a transformation of the input data then passing them through the network is the same as passing the input data through the network and then performing a transformation of the

output. CNNs are inherently translationally invariant; however, invariance does not extend straightforward to other types of transformations. G-CNNs [9,3,10] were introduced to tackle this issue by generalizing CNNs in a way such that symmetries are incorporated and fully exploited in the learning process. Very recently, PDE-G-CNNs [31,4] were proposed as PDEs-based framework based that generalized G-CNNs. The proposed PDEs were solved by providing analytical kernels approximations [31] and exact kernels sub-Riemannian approximations [4]. Intensive research on equivariant operators other than transformations is still conducted [29,20,33].

GANs [22,21] brought a new perspective to the deep learning community, deep learning with adversarial training is considered today as one of the most robust technique. With adversarial generative networks, there exists not only a good neural network-based classifier, referred to as the discriminator network, but also a generative network capable of producing realistic adversarial samples, all within a single architecture. This means that we now have a network that is aware of internal representations through its training to distinguish real inputs from artificial ones. Many extensions have been built for increasing its performances. Conditional GAN (CGAN) [19] was proposed as an extension of original GAN for generating facial images on the basis of facial attributes. Deep Convolutional GAN (DCGAN) [28] was proposed for image generation where both the generator and discriminator networks are convolutional. GRAN [24] is a GAN model based on a sequential process. Bidirectional GAN (BiGAN) and extensions [12,6] were proposed to map data into a latent code similar to an autoencoder. Generative Multi-Adversarial Network (GMAN) [16] was proposed for extending the minimax game to multiple players in GANs. In a different perspective, Wasserstein Generative Adversarial Network (WGAN) [1] was introduced to reduce the instability problems that occur during the training step, and also to eliminate the mode collapse effect. GANs and variants lack an inference mechanism.

In this work[4], we aim at providing noticeable improvements of former GAN models by using a geometric approach based on equivariant operators defined in a Lie group, and on mathematical morphology formulated in Riemannian manifolds. Main contributions can be summarized as follows: **1)** proposition of a new geometric generative model based on a new PDE-G-CNNs built on multiscale morphology operators and geometric image processing techniques, **2)** improvements of specific feature extraction while accounting intrinsic geometric features at multiple scales/levels, and **3)** equivariance of the network resulting in a complexity reduction and a geometric interpretability. Additional details and results that did not fit into the main paper can be found in supplementary material.

The paper is organized as follows. In Section 2, we define the notion of equivariance in Lie groups and present the group invariance property on Riemannian manifolds. In Section 3, we present the viscosity solutions for morphological dilations and erosions formulated as Lie group morphological convolutions in

[4] This work was partially supported by the ANR project Human4D ANR-19-CE23-0020

Riemannian manifolds. The proposed geometric generative (GM-GAN) model in presented in Section 4. Section 5 is dedicated to numerical experiments and comparisons with classical GAN models. The paper ends in Section 6 where concluding remarks and perspectives are discussed.

2 Equivariance and homogeneous spaces on Riemannian manifolds

Let M be a smooth manifold and $x \in M$. A linear mapping $v : C^\infty(M;\mathbb{R}) \to \mathbb{R}$ satisfying the Leibniz rule:

$$\forall\ f_1, f_2 \in C^\infty(M;\mathbb{R}) \quad v(f_1 f_2) = f_1(x)v(f_2) + v(f_1)f_2(x) \tag{1}$$

is called a derivation at x. For all $x \in M$, the set of derivations at x forms a real vector space of dimension d denoted T_xM so called the tangent space at x; its elements can be also called tangent vectors. In Euclidean space, an operator satisfying (1) is the derivative along a specific direction, and this definition is a generalization of derivatives on smooth manifolds in general.

Let G be a connected Lie group. We assume that the group G acts regularly on the spaces P and Q, meaning that there exists regular maps $\rho_P : G \times P \to P$ and $\rho_Q : G \times Q \to Q$ respectively defined for all $r, h \in G$, by:

$$\rho_P(rh, x) = \rho_P(r, \rho_P(h, x)) \text{ and } \rho_Q(rh, x) = \rho_Q(r, \rho_Q(h, x)), \tag{2}$$

making ρ_P and ρ_Q group actions on their respective spaces. In addition, we assume that the group G acts transitively on the spaces (smooth manifolds), meaning that for any two elements in these spaces, there exists a transformation in G that maps them to each other. This implies that P and Q can be viewed as homogeneous spaces.

Definition 1. *A Riemannian metric on a differentiable manifold M is given by a scalar product μ on each tangent space T_xM depending smoothly on the base point $x \in M$, that is, $\forall\ x \in M$, $\mu_x : T_xM \times T_xM \to \mathbb{R}$ is a symmetric, bilinear and positive definite map, and μ_x varies smoothly over M.*
A Riemannian manifold (M, μ) is a differentiable manifold M equipped with a Riemannian metric μ.

Definition 2. *Let G a connected Lie group with neutral element e and (M, μ) a connected Riemannian manifold. A left action of G on (M, μ) is an application $\varphi : G \times (M, \mu) \to (M, \mu)$ satisfying:*

1. $\varphi(e, x) = x$, $\forall\ x \in (M, \mu)$.
2. $\varphi(g, \varphi(h, x) = \varphi(gh, x)$, $\forall\ g, h \in G$ *and* $\forall\ x \in (M, \mu)$.

Let $\varphi : G \times (M, \mu) \to (M, \mu)$ be a left action of G on (M, μ). For a fixed $g \in G$, we define $\varphi_g : (M, \mu) \to (M, \mu)$; $x \mapsto \varphi_g(x) = \varphi(g, x)$.
The function $\varphi : G \times (M, \mu) \to (M, \mu)$ is a left action if $\forall\ g, h \in G$, one has:

$\varphi_e = id_M$ and $\varphi_g \circ \varphi_h = \varphi_{gh}$.
Let $\varphi_h : (M, \mu) \longrightarrow (M, \mu)$ be the left group action (considered here as a multiplication) by an element $h \in G$ defined $\forall\ x \in (M, \mu)$ by:

$$\varphi_h(x) = h \cdot x. \tag{3}$$

Let $\mathcal{L}_h$ be the left regular representation of G on functions f defined on M by $(\mathcal{L}_h f)(x) = f(\varphi_{h^{-1}}(x))$, with h^{-1} as the inverse of $h \in G$.

We consider a layer in a neural network as an operator (from functions on M_1 to functions on M_2). To ensure the equivarianc of the network, we shall require the operator to be equivariant with respect to the actions on the function spaces.

Let x_0 be an arbitrary fixed point on the connected Riemannian manifold (M, μ). Let $\pi : G \to (M, \mu)$ be the projection defined by assigning to each element h of G an element of (M, μ) in the following:

$$\forall\ h \in G \quad \pi(h) = \varphi_h(x_0). \tag{4}$$

In other words, once a reference point $x_0 \in (M, \mu)$ is chosen, the projection $\pi(h)$ assigns to every element h in G the unique point in (M, μ) to which h sends the chosen reference point x_0 under the action of φ_h given by (3).

In this work, we consider a connected Lie group G that acts transitively on the connected Riemannian manifold (M, μ). This means that for any points x and $y \in (M, \mu)$, there exists an element $h \in G$ such that $\varphi_h(x) = y$, corresponding to the definition of an homogeneous space under the action of the group G.

Definition 3. *Let G be a connected Lie group with homogeneous spaces M and N. Let ϕ be an operator on functions from M to functions on N. We say that ϕ is equivariant with respect to G if for all functions f, one has:*

$$\forall\ h \in G,\ (\phi \circ \mathcal{L}_h) f = (\mathcal{L}_h \circ \phi) f, \tag{5}$$

Let $h \in G$, $x \in (M, \mu)$ and $T_x M$ be the tangent space of (M, μ) at the point x. The pushforward of the group action φ_h denoted $(\varphi_h)_*$ is defined by: $(\varphi_h)_* : T_x M \to T_{\varphi_h(x)} M$ such that for all smooth functions f on (M, μ) and all $v \in T_x M$, one has: $((\varphi_h)_* v) f := v(f \circ (\varphi_h)_*)$.

For all $x \in (M, \mu)$, we refer to G-invariance of vector fields $X : x \mapsto T_x M$ if $\forall\ h \in G$ and for all differentiable functions f, one has $X(x)f = X(\varphi_h(x))[\mathcal{L}_h f]$.

Definition 4. *A vector field X on (M, μ) is invariant with respect to G if $\forall\ h \in G$ and $\forall\ x \in (M, \mu)$, one has: $X(\varphi_h(x)) = (\varphi_h)_* X(x)$.*

Definition 5. *A $(0, 2)$-tensor field μ on M is G-invariant if $\forall\ h \in G$, $\forall\ x \in M$ and $\forall\ v, w \in T_x(M)$, one has: $\mu|_h(v, w) = \mu|_{\varphi_h(x)}((\varphi_h)_* v, (\varphi_x)_* w)$.*

It follows from Definition 5 that properties derived from metric tensor field G invariance and vector field G invariance are the same.

Definition 6. *Let (M,μ) a connected Riemannian manifold, $x,y \in (M,\mu)$. The distance between x and y is defined as:* $d_\mu(x,y) = \inf_{\gamma \in \Gamma_t(x,y)} \int_0^t \sqrt{\mu|_{\gamma(t)}(\dot\gamma(t),\dot\gamma(t))}dt$, *with* $\Gamma_t(x,y) = \{\gamma : [0,t] \longrightarrow (M,\mu) \;\; \text{of class} \;\; C^1, \gamma(0) = x \;\; \text{and} \;\; \gamma(t) = y\}$.

Definition 7. *The cut locus is defined as the set of points $x \in M$ (or $h \in G$) from which the distance map is not smooth (except at x or h).*

Proposition 1. *Let $x,y \in (M,\mu)$ such that $\varphi_h(y)$ is away from the cut locus of $\varphi_h(x)$. Then, $\forall\ h \in G$, one has:* $d_\mu(x,y) = d_\mu\big(\varphi_h(x),\varphi_h(y)\big)$.

Remark 1. Staying away from the cut locus provides a unique distance in Definition 6. Also, thanks to Proposition 1, d_μ shares the same symmetries, since we derive it from a tensor field invariant under G.

3 Group morphological convolutions and PDEs

Link between morphological multiscale flat erosions and PDEs was established by running in $\mathbb{R}^n$ a first order Hamilton-Jacobi PDE type. Let (M,μ) be a compact and connected Riemannian manifold endowed with a metric μ, and $f, b : (M,\mu) \longrightarrow \mathbb{R}$.

Definition 8. *The group morphological convolution $\diamondsuit$ between b and f is defined $\forall\ x \in (M,\mu)$ by:* $b\diamondsuit f(x) = \inf_{p\in G}\{f(\varphi_p(x_0)) + b(\varphi_{p^{-1}}(x))\}$.

Denote TM the tangent bundle (M,μ) and $L : TM \to \mathbb{R}$ a Lagrangian function. Let $H : T^*M \to \mathbb{R}$ be the Hamiltonian associated to the Lagrangian L, H is defined on the cotangent bundle T^*M of (M,μ), $H(x,q) = \sup_{v\in T_xM}\{q(v) - L(x,v)\}$. The Hamilton-Jacobi PDE can be extended in Riemannian manifolds as follows: $\partial_t w + H\left(x, \nabla w\right) = 0$ in $(M,\mu)\times(0,+\infty)$; $w(\cdot,0) = f$ on (M,μ). Riemannian multiscale operations can be performed by choosing a specific Hamiltonian, respectively, $H = \|\nabla_\mu w\|_\mu^k$ for the multiscale dilations and $H = -\|\nabla_\mu w\|_\mu^k$ for multiscale erosions, and taking $k > 1$ allows to deal with more general structuring functions than the quadratic ones.

Proposition 2. *Let $f \in C^0((M,\mu),\mathbb{R})$ a continuous function and let $c_k = \frac{k-1}{k^{\frac{k}{k-1}}}$, $k > 1$. Viscosity solutions of the Cauchy problem:*

$$\frac{\partial w}{\partial t} + \|\nabla_\mu w\|_\mu^k = 0 \;\; in \;\; (M,\mu)\times(0;\ \infty);\ w(\cdot\,,0) = f \;\; on \;\; (M,\mu), \tag{6}$$

are given by: $f_t(x) = b_t^k \diamondsuit f(x) := \inf_{h\in G}\left\{ f\big(\varphi_h(x_0)\big) + c_k \frac{d_\mu\big(\varphi_{h^{-1}}(x), x_0\big)^{\frac{k}{k-1}}}{t^{\frac{1}{k-1}}} \right\}$,

where $b_t^k = c_k \frac{d_\mu(x_0,\cdot\,)^{\frac{k}{k-1}}}{t^{\frac{1}{k-1}}}$ *are the multiscale structuring functions.*

Proof. Viscosity solutions of the PDE (6) are given by HLO formulas [11]: $f_t(x) = \inf_{y\in M}\left\{f(y) + c_k\frac{d_\mu(x,y)^{\frac{k}{k-1}}}{t^{\frac{1}{k-1}}}\right\}$. The projection π (4) is defined by associating any $h \in G$ to an element $x \in (M,\mu)$. Then, using the definition and accounting the invariance property in Proposition 1, one gets:

$$\begin{aligned} f_t(x) &= \inf_{h\in G}\left\{f\big(\varphi_h(x_0)\big) + c_k\frac{d_\mu(x,\varphi_h(x_0))^{\frac{k}{k-1}}}{t^{\frac{1}{k-1}}}\right\} \\ &= \inf_{h\in G}\left\{f\big(\varphi_h(x_0)\big) + c_k\frac{d_\mu\big(\varphi_{h^{-1}}(x),x_0\big)^{\frac{k}{k-1}}}{t^{\frac{1}{k-1}}}\right\} \\ &= \inf_{h\in G}\left\{f\big(\varphi_h(x_0)\big) + b_t^k\big(\varphi_{h^{-1}}(x)\big)\right\} = b_t^k \diamondsuit f(x). \square \end{aligned}$$

By reversing the time, we can prove that the viscosity solutions of the Cauchy problem corresponding to multiscale dilations:

$$\frac{\partial w}{\partial t} - \|\nabla_\mu w\|_\mu^k = 0 \text{ in } (M,\mu)\times(0;\ \infty);\ w(\cdot\,,0) = f \text{ on } (M,\mu) \tag{7}$$

are given by [11]: $f^t(x) = \sup_{x\in(M,\mu)}\left\{f(y) - C_k\frac{d_\mu(x,y)^{\frac{k}{k-1}}}{t^{\frac{1}{k-1}}}\right\}$, and thus, using the same arguments as in the preceding proof, one has: $f_t(x) = -(b_t^k\diamondsuit(-f))(x)$.

Proposition 3. *Let $k > 1$. For all $t, s \geq 0$, the family of structuring functions b_t^k satisfy the following semigroup property: $b_{t+s}^k = b_t^k \diamondsuit b_s^k$.*

4 Morphological equivariant PDEs for generative models

We aim at proposing generative models for images that are based on PDEs satisfying an equivariance property. Our approach is resumed in two major steps: **1)** design of morphological PDEs in Riemannian manifolds akin to Section 3 as alternatives for introducing non-linearities in traditional CNNs that preserve an equivariant processing in the composition of the feature maps in layers, and **2)** proposition of a generative model based on this structure and classical GANs.

4.1 Morphological PDE-based layers

Feature maps are carried out in traditional CNNs throughout the classical convolution, pooling and ReLU activation functions. Our goal is to propose PDEs that behave like traditional CNNs, in one hand, and preserve an equivariance property, on the other hand. For that purpose, PDEs will be formulated on group transformations to ensure equivariance and make PDEs consistent with G-CNNs [9,3,10]. Equivariance is a robust way to incorporate desired and essential symmetries into the network so that there is no more need to learn such symmetries;

consequently, the amount of data is reduced. Viewing layers as image processing operators allows us use well elaborated image analysis and processing techniques to design the network. Thin image analysis is needed to achieve our objective. Due to its nonlinearity aspects, good shape and geometry description capabilities, mathematical morphology appeared as an efficient and powerful tool for multiscale image and data analysis [30]. For a better analysis of geometrical image structures, it is also interesting to consider works from geometric image analysis [34,17,13,5,15]. Image and data analysis and processing methods based on non-Euclidean metrics; for instance, Riemannian metrics, are well known to improve a lot Euclidean based approaches. Riemannian manifolds are proved to behave very well for capturing thin data structures, providing then better representations and analysis of geometrical structures present in the data. This fact is shown in many image processing studies with real life applications; for instance, in video surveillance, shape and surface analysis, human body and face analysis, image segmentation [32,2,7,26,35,27]. For these reasons, we choose homogeneous spaces to avoid Euclidean metrics so that the network is provided with image processing capabilities for a better handling of geometric thin structures [8,25,18,14,11]. Doing so should make feature maps richer, and combined with the equivariance property of the morphological PDEs will provide neat improvements of classical GANs in terms of quality of the content generation. Morphological PDEs are thus used to replace the pooling operations and ReLU activation functions in the proposed generative model.

4.2 PDE model design

PDE-G-CNNs were formally introduced in homogeneous spaces with G-invariance metric tensor fields on quotient spaces [31]. Built on the primary approach, the proposed model is based on a combination of traditional CNNs and morphological PDE layers of Hamilton-Jacobi type in Riemannian manifolds, and is composed of the following PDEs:

- Convection: $\dfrac{\partial w}{\partial t} + \alpha w = 0$ in $(\mathcal{M}, \mu) \times (0,\ \infty)$; $w(\cdot, 0) = f$ on $(\mathcal{M}, \mu)$.
- Diffusion: $\dfrac{\partial w}{\partial t} + (-\Delta_\mu)^{k/2} w = 0$ in $(\mathcal{M}, \mu) \times (0,\ \infty)$; $w(\cdot, 0) = f$ on $(\mathcal{M}, \mu)$.
- Morphological multiscale erosions and dilations for $(+)$ and $(-)$ sign:

$$\frac{\partial w}{\partial t} \pm \|\nabla_\mu w\|_\mu^k = 0 \text{ in } (\mathcal{M}, \mu) \times (0,\ \infty);\ w(\cdot, 0) = f \text{ on } (\mathcal{M}, \mu), \tag{8}$$

where α a is vector field invariant under G on $(\mathcal{M}, \mu)$, Δ_μ represents the Laplace-Beltrami operator, $\|\cdot\|_\mu$ the norm induced by the Riemannian metric μ and $k > 1$. The above system of PDEs consitutes the PDE model solved in a step basis using the operator splitting method, where each step corresponds to one of the PDEs. In this work, we only use the morphological multiscale operations steps (8), the convection and diffusion terms are left for future work. PDEs (8) introduce nonlinearities into the generator network of the GM-GAN using morphological convolutions, which are obtained a viscosity sense and given respectively for multiscale dilations and erosions thanks to Proposition 2.

Proposition 4. *Let $f \in C^{\infty}((\mathcal{M},\mu))$ and $B \subset (\mathcal{M},\mu)$ an non-empty set. Consider the flat structuring function $b : (\mathcal{M},\mu) \to \mathbb{R} \cup \{\infty\}$. Then, one has:* $-\left(b\diamondsuit(-f)\right)(x) = \sup_{\substack{h\in G \\ \varphi_{h^{-1}}(x)\in B}} f\left(\varphi_h(x_0)\right)$.

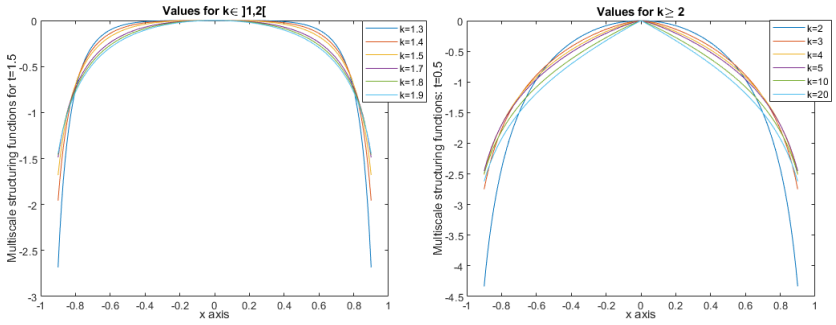
The max pooling of function f with motif B can in fact be seen as a flat morphological dilation with a structurant element B. It is truly the case for example for $\mathbb{R}^n$. Indeed, for $f \in C^0\left(\mathbb{R}^n\right)$ and $B \subset \mathbb{R}^n$ a compact set, for every $x \in \mathbb{R}^n$, one has:$-\left(b\diamondsuit_{\mathbb{R}^n}(-f)\right)(x) = \sup_{y\in B} f(x-y)$, where the right hand side is in fact a flat dilation with a structurant element B.

Proposition 5. *Let $f \in C_c^0((\mathcal{M},\mu))$. Morphological dilation with the following structuring function: $b(x) = 0$, if $x = x_0$; and $b(x) = \sup_{x\in\mathcal{M}} f(x)$, otherwise, is exactly the same as applying a ReLU to f:* $-\left(b\diamondsuit(-f)\right)(x) = \max\{0, f(x)\}$.

4.3 Architecture of morphological equivariant PDEs based on GAN

Similarly to GAN, the proposed geometric morphological GAN (GM-GAN) is composed of two networks: a generator (G) and a discriminator (D) which are both multi-layer perceptrons. As detailed in the preceding section, we introduce into the network G morphological PDE-based layers through the resolution in a step basis of Hamilton-Jacobi PDEs (8), whose viscosity solutions are given for multiscale erosions and dilations thanks to Proposition 2. To deal with computation issues and practical implementation of the proposed framework, we take advantage of the geometric properties of hyperbolic spaces and generate various and rich content on data with multiple transformations. For doing so, we provide the distance d_μ in the geodesic ball by considering the hyperbolic ball $B = \{(x_1, x_2) \in \mathbb{R}^2 \text{ such that } x_1^2 + x_2^2 < 1\}$, which is endowed with the metric $\mu = \dfrac{4(\mathrm{d}x_1^2 + \mathrm{d}x_2^2)}{(1-\|x\|^2)^2}$, where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^2$. The distance is obtained as follows: $d_\mu(x,y) = \text{Argcosh}\left(1 + \dfrac{2\|x-y\|^2}{(1-\|x\|^2)(1-\|y\|^2)}\right)$.
Concave structuring functions $b_t^k = c_k \dfrac{d_\mu(x_0,\cdot)^{\frac{k}{k-1}}}{t^{\frac{1}{k-1}}}$ are represented in Fig. 1 for different values of t and k in $]-1;\ 1[$.

GM-GAN training procedure remains the same as in traditional GANs. Specifically, the training procedure is carried out separately but simultaneously. The model takes as input some noise z defined with a prior probability p_z, and then, attempts to learn the distribution of the generator p_g, by representing a function $G(z;\theta_g)$ from z to the data space. The discriminator network D takes an input image x and finds a function $D(x;\theta_d)$ from x to a single scalar, which is the probability that the image x comes from p_{data} which defines the origin of the sampled images. The output of the D network returns a value close to 1 if x is a real image from p_{data}, and a value very close to 0 if x comes from p_g; otherwise.



(a) (b)

Fig. 1: $b_t^k(x)$, $x \in]-1;\ 1[$: (a) for $t = 1.5$ and $k \in]1;\ 2[$. (b) for $t = 0.5$ and $k \geq 2$.

The main objective of network D is to maximize $D(x)$ for an image coming from the true data distribution p_{data}, while minimizing $D(x) = D(G(z;\theta_g))$ for images generated from p_z and not from p_{data}. The objective of the generator G is to deceive the D network, meaning to maximize $D(G(z;\theta_g))$. This is equivalent to minimize $1 - D(G(z;\theta_g))$ as D is a binary classifier. This conflict between these objectives is called the minimax game and formulated as follows: $\min\max E_{x\sim p_{data}(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1 - D(G(z;\theta_g)))]$. The case $p_g = p_{data}$ corresponds to the global optimum of the minimax game. Main contributions of the proposed GM-GAN rely on the equivariance property and non linearity characteristics brought out by group morphological convolutions and their ability to extract thin geometrical features, which lead to richer feature maps and a reduction of the amount training data.

For the GM-GAN generator, let x be the input data into the morphological layer called *Morphoblock*. Then, x goes first through a multiscale morphological erosion operation, followed by a multiscale morphological dilation. Afterwards, both erosion and dilation are followed by a linear convolution. The output of the PDE layer is obtained by a linear combination of the two outputs. The overall architecture of the GM-GAN generator is illustrated in Fig. 2.

5 Numerical experiments

GM-GAN and GAN are applied to MNIST dataset. MNIST database consists of $70,000$ black-and-white 28x28 images that represent handwritten digits from 0 to 9. It is divided into a training set of $60,000$ images and a test set of $10,000$ images. Same training parameters are set for GM-GAN and GAN: number of epochs to 200, the batch size to 64, the latent space dimensionality to 100, and the interval between image samples to 400. Generated images with GM-GAN and GAN are displayed in Fig. 3 showing higher generation quality with GM-GAN in comparison to traditional GAN. This can be seen by comparing images

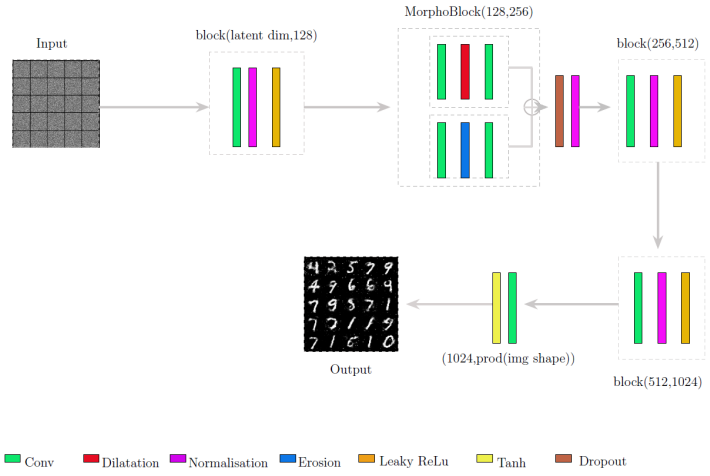
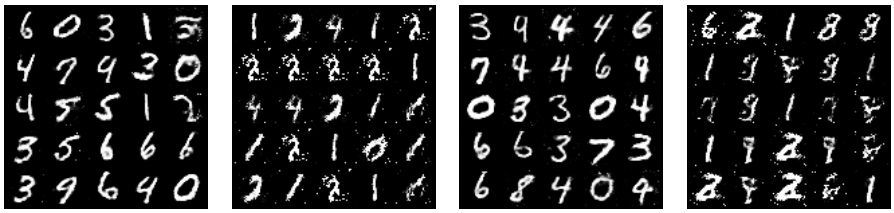


Fig. 2: Architecture of GM-GAN generator.

produced at epochs 70 to 95 with GM-GAN (Figs. 3a, 3e, 3i, 3m and 3q) and those generated with GAN at same epochs (Figs. 3b, 3f, 3j, 3n and 3r). For instance, some digits are clearly identifiable with GM-GAN based generation, whereas it is almost impossible to recognize the digits with GAN based ones. We also observe that the images generated with GM-GAN at epochs going from 100 to 120 (Figs. 3c, 3g, 3k, 3o and 3s) are of better quality than generated ones with GAN for the last five epochs going from epoch 195 to 199 (Figs. 3d, 3h, 3l, 3p and 3t). To better discriminate that fact, we zoom in on some areas in images generated at epochs 85, 92 and 96 (Figs. 4-(a)-(b), (c)-(d) and (e)-(f); respectively), and highlight the realistic variations between the generated images of the same digit. This indicates that GM-GAN has a deeper understanding of the sample characteristics and is capable of generalizing them beyond the specific examples they are trained on. This can be observed in Fig. 4-(b) with digits 3 and 6, in Fig. 4-(d) with digits 2 and 8, and in Fig. 4-(f) with digits 9 and 7.

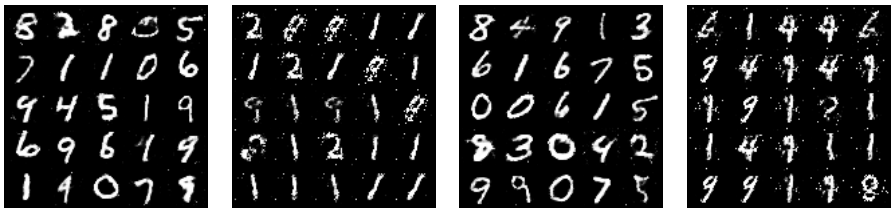
GM-GAN complexity is also reduced throughout the equivariance property by eliminating the need to learn symmetries. This is illustrated by reducing MNIST training dataset by a half and comparing generated images at epoch 42. GM-GAN results (Fig. 5a) show again better image quality and high variations of generated digits in comparison to GAN (Fig. 5b). Results highlight the importance of equivariance in morphological operators, turning out to dataset reduction without significantly impacting generation results (see Fig. 5c for GM-GAN and Fig. 5d for images generated at the same epoch using the hole dataset).

To highlight again the usefulness of morphological equivariant operators, we apply both GM-GAN and GAN models on RotoMNIST; generated images are displayed in Fig. 6. It can be seen in results obtained with GM-GAN from epoch

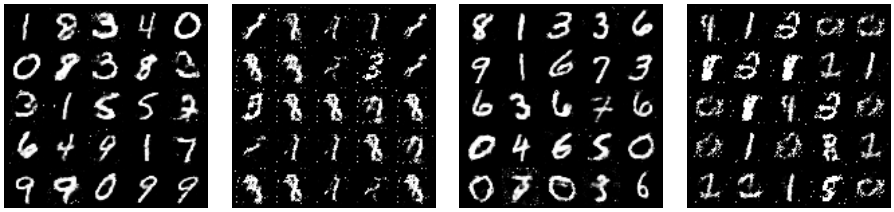


(a) GM-GAN: 75 (b) GAN: 75 (c) GM-GAN: 100 (d) GAN: 195

(e) GM-GAN: 80 (f) GAN: 80 (g) GM-GAN: 105 (h) GAN: 196



(i) GM-GAN: 85 (j) GAN: 85 (k) GM-GAN: 110 (l) GAN: 197

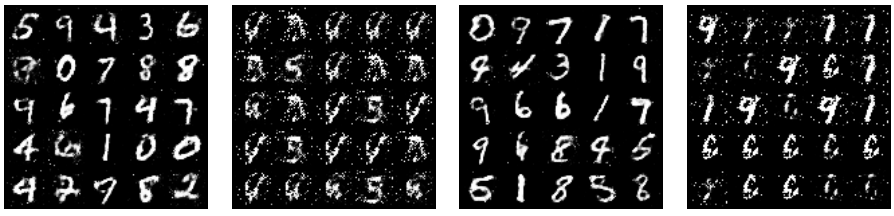


(m) GM-GAN:90 (n) GAN: 90 (o) GM-GAN: 115 (p) GAN: 198

(q) GM-GAN: 95 (r) GAN: 95 (s) GM-GAN: 120 (t) GAN: 199

Fig. 3: Image generation using MNIST: GM-GAN vs. GAN.

Fig. 4: Zoom in on images generated with GM-GAN at different epochs.



(a) GM-GAN(1/2) (b) GAN(1/2) (c) GM-GAN (d) GAN

Fig. 5: GM-GAN vs. GAN at epoch 42 with half (1/2) and whole MNIST dataset.

70 to 95 (Figs. 6a, 6e, 6i, 6m, and 6q) that digits are clearly identifiable and far better than those generated with GAN at the same epochs (Figs. 6b, 6f, 6j, 6n, and 6r) where digits are barely formed. The same is noticed with GM-GAN from epoch 100 to 120 (Figs. 6c, 6g, 6k, 6o, and 6s), in comparison with GAN for the last 5 epochs (Figs. 6d, 6h, 6l, 6p, and 6t). This demonstrates that GM-GAN is more suitable for data under rotation transformations, and highlights one more time the importance of equivariance for generating satisfactory results under various transformations.

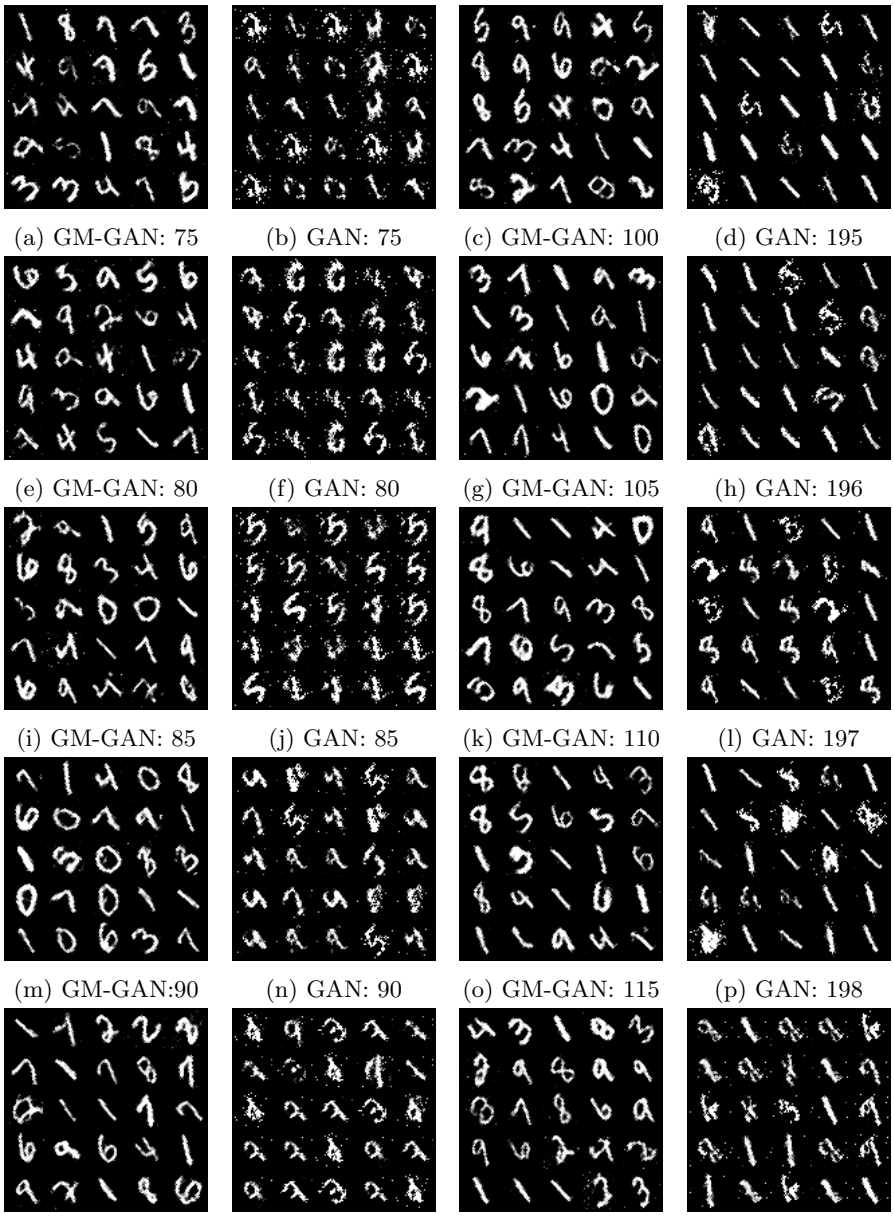
Quantitative evaluations are provided using the Fréchet Inception Distance (FID). A low FID indicates a high similarity between generated and real data, corresponding to good generation quality. In Fig. 7, we present the FID curves of both models over epochs (taking FID of generated images at intervals of 10 epochs) on both MNIST and RotoMNIST datasets. It can be seen that starting from epoch 40, FIDs of GM-GAN generated results are significantly lower than ones generated using GAN, which confirms the qualitative results discussed just above.

6 Conclusion and perspectives

We have proposed here a geometric generative GM-GAN model based on PDE-G-CNNs and built from derived equivariant morphological operators and geometric image processing techniques. The proposed equivariant morphological PDE layers are composed of multiscale dilations and erosions without any need to approximate convolutions kernels, and meanwhile, group symmetries are defined on Lie groups allowing a geometrical interpretability of GM-GAN with left invariance properties. As shown by preliminary results on MNIST and RotoMNIST datasets, preliminary qualitative and quantitative results show noticeable improvements compared classical GAN. Indeed, thin image features are better extracted by accounting intrinsic geometric features at multiscale levels, and the network complexity is reduced. The proposed approach can be extended to various generative models. Future works include applying GM-GAN on other datasets, designing fully equivariant generative models entirely based on PDE-G-CNNs, and studying GM-GAN complexity to demonstrate the computational advantages of the proposed model over classical GAN.

References

1. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: International conference on machine learning. pp. 214–223. PMLR (2017)
2. Balan, V., Stojanov, J.: Finslerian-type GAF extensions of the riemannian framework in digital image processing. Filomat **29**(3), 535–543 (2015)
3. Bekkers, E.J., Lafarge, M.W., Veta, M., Eppenhof, K.A., Pluim, J.P., Duits, R.: Roto-translation covariant convolutional networks for medical image analysis. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018: 21st International Conference, Proceedings, Part I. pp. 440–448. Granada, Spain (Sep 2018)



(a) GM-GAN: 75 (b) GAN: 75 (c) GM-GAN: 100 (d) GAN: 195

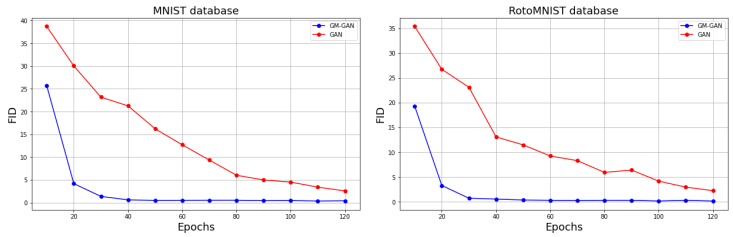
(e) GM-GAN: 80 (f) GAN: 80 (g) GM-GAN: 105 (h) GAN: 196

(i) GM-GAN: 85 (j) GAN: 85 (k) GM-GAN: 110 (l) GAN: 197

(m) GM-GAN:90 (n) GAN: 90 (o) GM-GAN: 115 (p) GAN: 198

(q) GM-GAN: 95 (r) GAN: 95 (s) GM-GAN: 120 (t) GAN: 199

Fig. 6: Image generation using RotoMNIST: GM-GAN vs. GAN.



(a) (b)

Fig. 7: FID using GM-GAN vs. GAN with: (a) MNIST. (b) RotoMNIST.

4. Bellaard, G., Bon, D.L., Pai, G., Smets, B.M., Duits, R.: Analysis of (sub-)Riemannian PDE-G-CNNs. Journal of Mathematical Imaging and Vision pp. 1–25 (2023)
5. Burger, M., Sawatzky, A., Steidl, G.: First order algorithms in variational image processing. Springer (2016)
6. Chen, M., Denoyer, L.: Multi-view generative adversarial networks. In: Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2017, Skopje, Macedonia, September 18–22, 2017, Proceedings, Part II 10. pp. 175–188. Springer (2017)
7. Citti, G., Franceschiello, B., Sanguinetti, G., Sarti, A.: Sub-riemannian mean curvature flow for image processing. SIAM Journal on Imaging Sciences **9**(1), 212–237 (jan 2016)
8. Citti, G., Sarti, A.: A cortical based model of perceptual completion in the roto-translation space. Journal of Mathematical Imaging and Vision **24**, 307–326 (2006)
9. Cohen, T., Welling, M.: Group Equivariant Convolutional Networks. In: International conference on machine learning. pp. 2990–2999. PMLR (2016)
10. Cohen, T.S., Geiger, M., Weiler, M.: A general theory of equivariant cnns on homogeneous spaces. Advances in neural information processing systems **32** (2019)
11. Diop, E.H.S., Mbengue, A., Manga, B., Seck, D.: Extension of Mathematical Morphology in Riemannian Spaces. In: Lecture Notes in Computer Science, pp. 100–111. Springer International Publishing (2021)
12. Donahue, J., Krähenbühl, P., Darrell, T.: Adversarial feature learning. arXiv preprint arXiv:1605.09782 (2016)
13. Dubrovina-Karni, A., Rosman, G., Kimmel, R.: Multi-region active contours with a single level set function. IEEE transactions on pattern analysis and machine intelligence **37**(8), 1585–1601 (2014)
14. Duits, R., Bekkers, E.J., Mashtakov, A.: Fourier transform on the homogeneous space of 3D positions and orientations for exact solutions to linear PDEs. Entropy **21**(1), 38 (2019)
15. Duits, R., Burgeth, B.: Scale spaces on Lie groups. In: International Conference on Scale Space and Variational Methods in Computer Vision. pp. 300–312 (2007)
16. Durugkar, I., Gemp, I., Mahadevan, S.: Generative multi-adversarial networks. arXiv preprint arXiv:1611.01673 (2016)
17. Fadili, J., Kutyniok, G., Peyré, G., Plonka-Hoch, G., Steidl, G.: Guest editorial: mathematics and image analysis. Journal of Mathematical Imaging and Vision **52**, 315–316 (2015)

18. Franceschiello, B., Mashtakov, A., Citti, G., Sarti, A.: Geometrical optical illusion via sub-Riemannian geodesics in the roto-translation group. Differential Geometry and its Applications **65**, 55–77 (2019)
19. Gauthier, J.: Conditional generative adversarial nets for convolutional face generation. Class project for Stanford CS231N: convolutional neural networks for visual recognition, Winter semester **2014**(5), 2 (2014)
20. Gerken, J.E., Aronsson, J., Carlsson, O., Linander, H., Ohlsson, F., Petersson, C., Persson, D.: Geometric deep learning and equivariant neural networks. Artificial Intelligence Review **56**(12), 14605–14662 (Jun 2023)
21. Goodfellow, I.: Generative Adversarial Networks. In: NIPS. p. 57 (2017)
22. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. Advances in neural information processing systems **27** (2014)
23. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., TingLiu, Wang, X., Wang, L., Wang, G., Cai, J., Chen, T.: Recent Advances in Convolutional Neural Networks. Pattern recognition **77**, 354–377 (2018)
24. Im, D.J., Kim, C.D., Jiang, H., Memisevic, R.: Generating images with recurrent adversarial networks. arXiv preprint arXiv:1602.05110 (2016)
25. Janssen, M.H., Janssen, A.J., Bekkers, E.J., Bescós, J.O., Duits, R.: Design and processing of invertible orientation scores of 3D images. Journal of mathematical imaging and vision **60**, 1427–1458 (2018)
26. Kurtek, S., Jermyn, I.H., Xie, Q., Klassen, E., Laga, H.: Elastic shape analysis of surfaces and images. In: Riemannian Computing in Computer Vision, pp. 257–277. Springer International Publishing (2016)
27. Pierson, E., Daoudi, M., Tumpach, A.B.: A Riemannian Framework for Analysis of Human Body Surface. In: IEEE/CVF Winter Conference on Applications of Computer Vision, WACV. pp. 2763–2772. Waikoloa, HI, USA (Jan 2022)
28. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015)
29. Romero, D., Bekkers, E., Tomczak, J., Hoogendoorn, M.: Attentive Group Equivariant Convolutional Networks. In: Proceedings of Machine Learning Research. pp. 8188–8199 (2020)
30. Shih, F.: Image processing and mathematical morphology : fundamentals and applications. CRC Press, Boca Raton (2009)
31. Smets, B.M.N., Portegies, J., Bekkers, E.J., Duits, R.: PDE-Based Group Equivariant Convolutional Neural Networks. Journal of Mathematical Imaging and Vision **65**(1), 209–239 (2022)
32. Su, J., Kurtek, S., Klassen, E., Srivastava, A.: Statistical analysis of trajectories on Riemannian manifolds: Bird migration, hurricane tracking and video surveillance. The Annals of Applied Statistics **8**(1) (Mar 2014)
33. Tian, C., Zhang, Y., Zuo, W., Lin, C.W., Zhang, D., Yuan, Y.: A Heterogeneous Group CNN for Image Super-Resolution. IEEE Transactions on Neural Networks and Learning Systems pp. 1–13 (2024)
34. Welk, M., Weickert, J.: Pde evolutions for m-smoothers: from common myths to robust numerics. In: International Conference on Scale Space and Variational Methods in Computer Vision. pp. 236–248. Springer (2019)
35. Younes, L.: Shapes and Diffeomorphisms. Springer Berlin Heidelberg (May 2019)