RELATIVE STABILITY TOWARD DIFFEOMORPHISMS INDICATES PERFORMANCE IN DEEP NETS

Leonardo Petrini, Alessandro Favero,
Mario Geiger, Matthieu Wyart

EPFL — NEURAL INFORMATION PROCESSING SYSTEMS

IN SHORT

1. Why **deep nets** can classify images living in **high dimension?**
2. Do they do so by becoming **stable to diffeomorphisms?**
3. Previous empirical works are in support of a negative answer to (2.)

We revisit this question by

4. defining a **maximum-entropy distribution on diffeomorphisms**
5. finding that the ***stability toward diffeomorphisms relative to that of generic transformations*** R_f shows robust empirical behaviors across architectures and datasets.

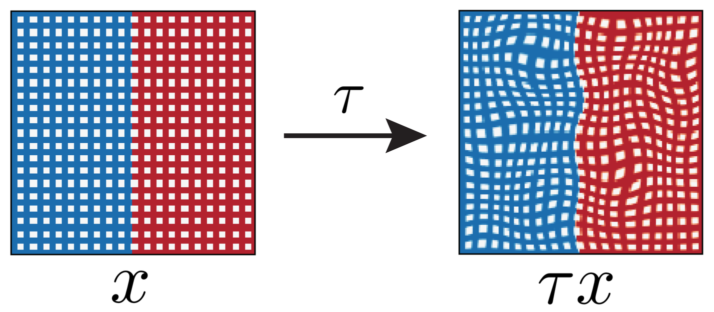
We conclude that ***relative diffeomorphisms stability***

6. **builds up with training** in SOTA nets
7. is important for obtaining good **performance**

MAX-ENTROPY DIFFEOMORPHISMS

Goal: define a distribution of *typical* diffeomorphisms of controlled magnitude.

NOTATION	
$x(s)$	input image intensity
$s=(u,v)\in[0,1]^2$	(continuous) pixel position
τ	smooth deformation
$[\tau x](s)=x(s-\tau(s))$	deformed image
$\tau(s)=(\tau_u(s),\tau_v(s))$	displacement vector field



The **deformation amplitude** is measured by:

$$\|\nabla\tau\|^2=\int_{[0,1]^2}(\nabla\tau_u)^2+(\nabla\tau_v)^2\,dudv. \qquad (1)$$

We write each of the displacement fields in real Fourier basis and fix the picture frame not to be deformed (i.e. $\tau_u=\tau_v=0$ if $u=0,1$ or $v=0,1$):

$$\tau_u=\sum_{i,j}C_{ij}\sin(i\pi u)\sin(j\pi v). \qquad (2)$$

The Fourier coefficients C_{ij} are random variables and their distribution defines the one over displacement fields.

To determine it, we inject (2) into (1) and get:

$$\|\nabla\tau\|^2=\frac{\pi^2}{4}\sum_{i,j}(C_{ij}^2+D_{ij}^2)(i^2+j^2).$$

In order to control $\langle\|\nabla\tau\|^2\rangle$, we introduce a **temperature** parameter T such that $\langle\|\nabla\tau\|^2\rangle\propto T$. This is a *canonical ensemble* where the energy $\|\nabla\tau\|^2$ is a sum of quadratic random variables and the *equipartition theorem* applies: each term contributes equally to the sum -on average- fixing the variance of the Fourier coefficients:

$$\langle C_{ij}^2\rangle=\frac{T}{i^2+j^2}.$$

Finally, the distribution that **maximizes entropy** with a constraint on the variance is the **Gaussian**, hence

$$C_{ij}\sim\mathcal{N}\left(0,\frac{T}{i^2+j^2}\right).$$

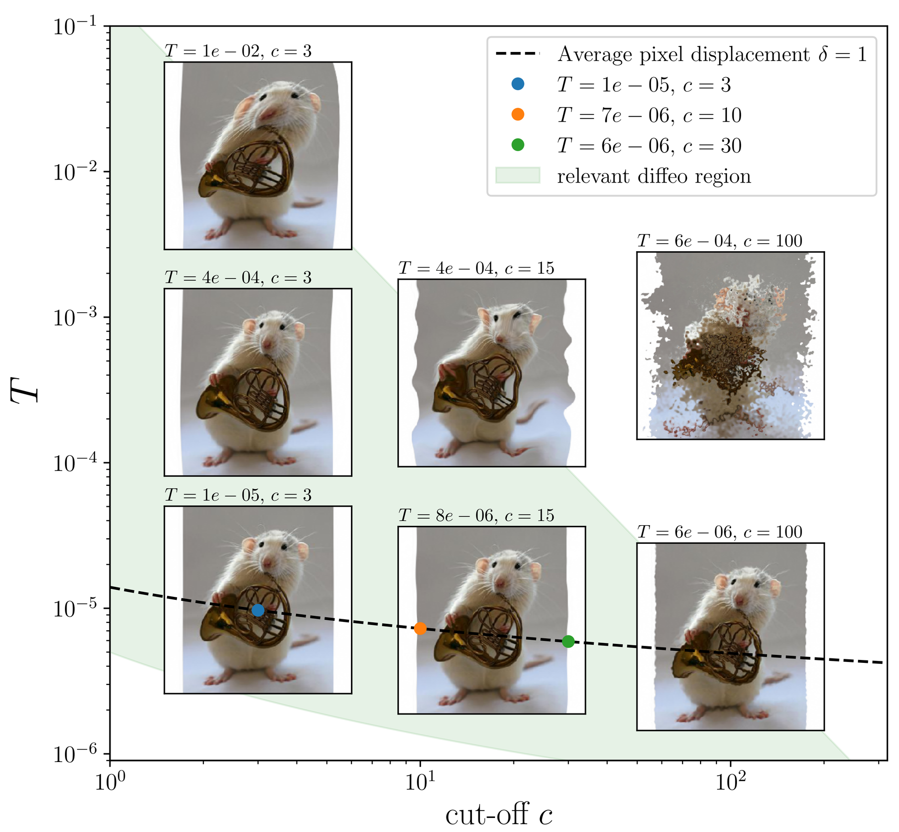


Figure: some sample of ***max-entropy diffeomorphisms***.

RELATIVE STABILITY

Goal: establish if a **deep net** learns to become more stable to **diffeomorphisms** than to a **generic data transformation**.

NOTATION	
x	input image intensity
τ	smooth deformation
η	isotropic noise
f	network function

Moreover, we fix $\|\eta\|=\langle\|\tau x-x\|\rangle$

We define the **relative stability to diffeomorphisms** as

$$R_f=\frac{\langle\|f(\tau x)-f(x)\|^2\rangle_{x,\tau}}{\langle\|f(x+\eta)-f(x)\|^2\rangle_{x,\eta}}.$$

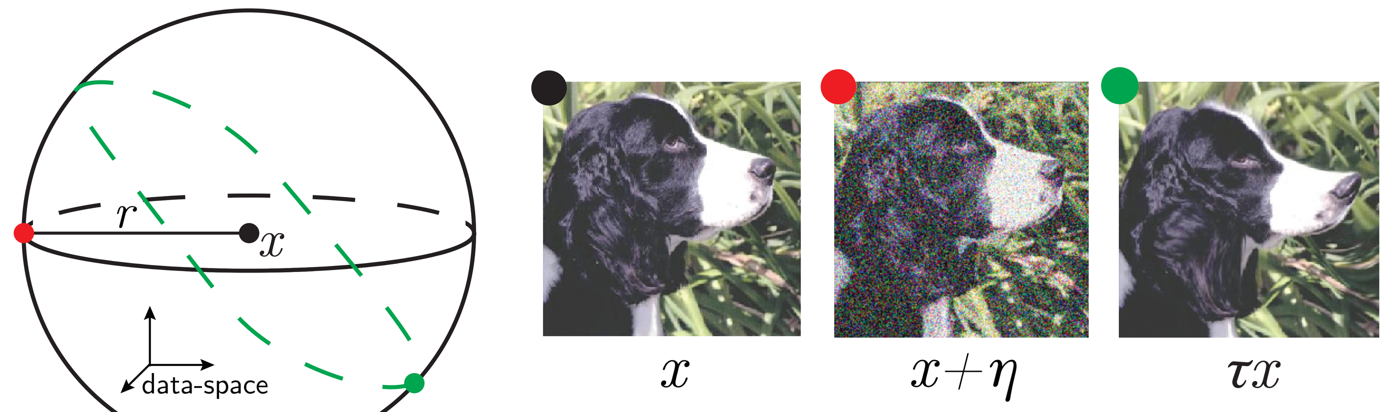
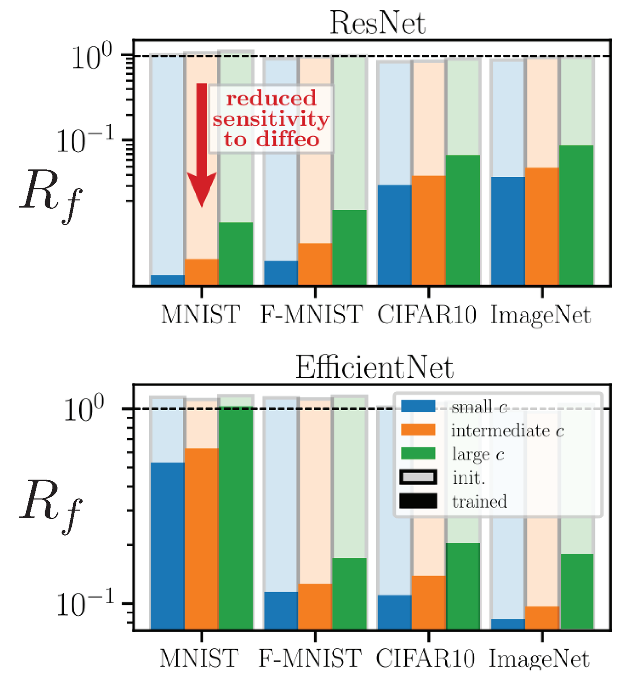
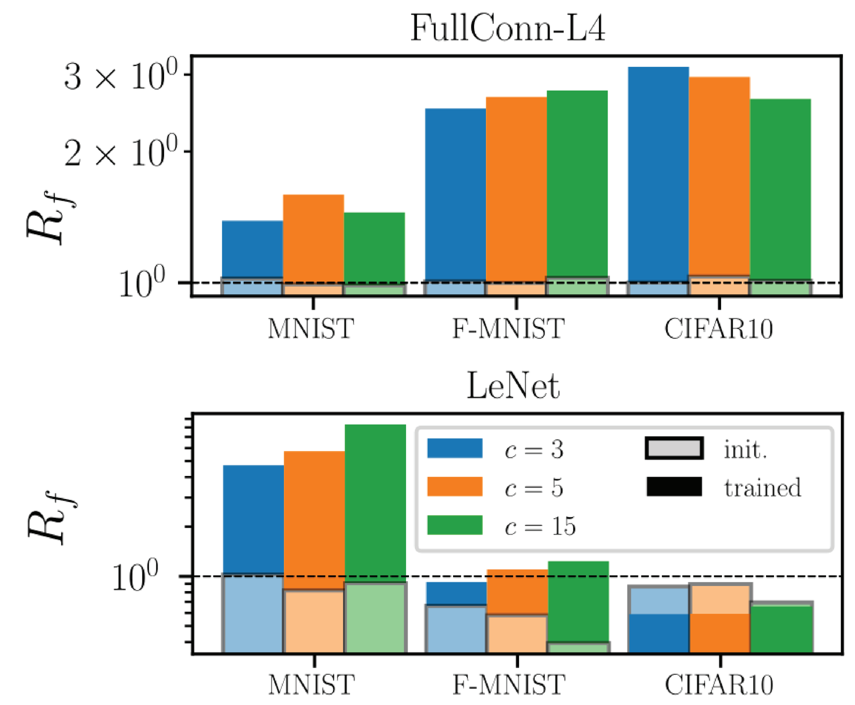


Figure: data-space around a data-point (illustrative). The relative stability to diffeomorphisms characterizes how a net f varies in the green directions, normalized by random ones - i.e. any point on the sphere, e.g. red dot.

DEEP NETS LEARN DIFFEO INVARIANCE

1. At **initialization** (shaded bars) $R_f\approx1$: neural nets, at initialization, are not particularly stable to diffeomorphisms.
2. **After training** (full bars) R_f is reduced by one or two orders of magnitude for SOTA nets, but not for simple ones.

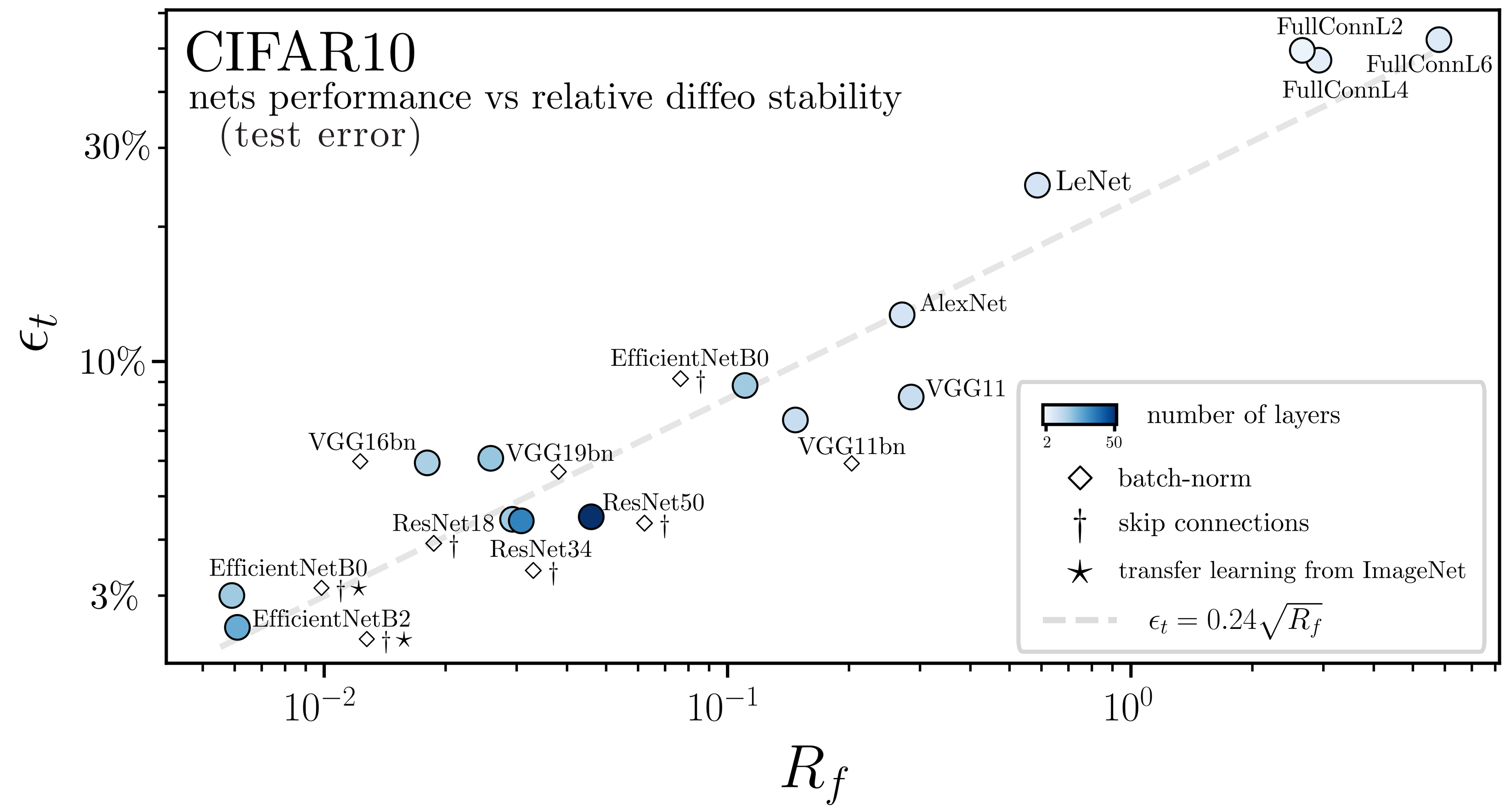




Is there a relationship with performance?

DIFFEO STABILITY AND PERFORMANCE

Relative stability shows a striking correlation with performance!



CONCLUSION

We have introduced a novel empirical framework to characterize how **deep nets become invariant to diffeomorphisms**. It is jointly based on a maximum-entropy distribution for diffeomorphisms, and on the realization that stability of these transformations relative to generic ones R_f strongly correlates to performance, instead of just the diffeomorphisms stability considered in the past.

Diffeomorphisms library: github.com/pcsl-epfl/diffeomorphism
Training SOTA nets: github.com/leonardopetrini/diffeo-sota
Pre-trained models:: doi.org/10.5281/zenodo.5589869

leonardopetrini@epfl.ch
github.com/leonardopetrini
twitter.com/leopetrini_