Mathematical models for the visual pathway and Deep Learning

Rita Fioresi, FaBiT, Unibo

Unibo, 2024

May 28, 2024

Join CaLISTA CA 21109!

https://site.unibo.it/calista/en

- Working group 1: Cartan Geometry and Representation Theory
- Working group 2: Integrable Systems and Supersymmetry
- Working group 3: Noncommutative Geometry and Quantum Homogeneous Spaces
- Working group 4: Vision
- Working group 5: Dissemination and Public Engagement

https://e-services.cost.eu/action/CA21109/working-groups/applications

Photographic Contest

A photographic contest on the theme of 'Symmetry in mathematics and physics'

- A picture on 'Symmetry in Mathematics and Physics' should be sent to the CaLISTA email address calistaeuproject@gmail.com
- Short essay explaining why the picture is related 'Symmetry in Mathematics and Physics'.
- The prize is a maximum of 500 Euros of expenses reimbursement towards the participation in the CaLISTA Workshop 'Geometry Informed Machine Learning taking place the 2-5 September, 2024 in Paris.
- The opening date is 1st of April and the closing date is 31st of May. The decision will be made by mid June.

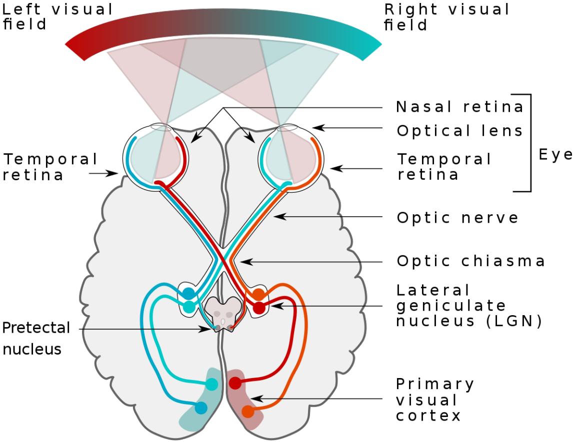
Plan of the Talk

1. The human visual pathway
2. Mathematical modeling of the visual cortex
3. A sub-Riemannian approach to the question of border completion
4. A Deep Learning approach to border enhancement and light sensitivity

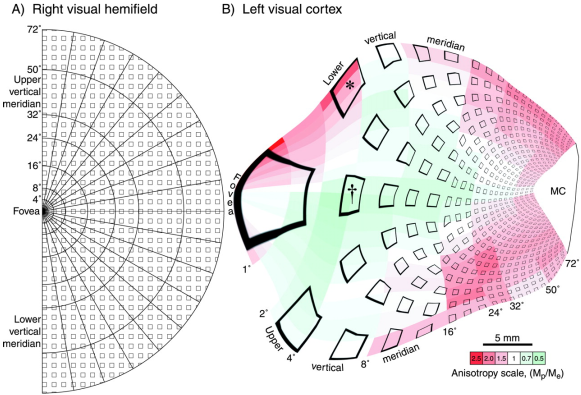
Bibliography

- R. Fioresi, A. Marraffa, J. Petkovic, A new perspective on border completion in visual cortex as bicycle rear wheel geodesics paths via sub Riemannian Hamiltonian formalism, Differential Geometry and its Applications, 10.1016/j.difgeo.2024.102125, 2024.
- J. Petkovic, R. Fioresi, A precortical module for robust CNNs to light variations, to appear in Neural Computations, 2024.
- L. Grementieri, R. Fioresi Model-centric Data Manifold: the Data Through the Eyes of the Model SIAM J. Imaging Sci. 15 (2022), no. 3, 1140–1156.

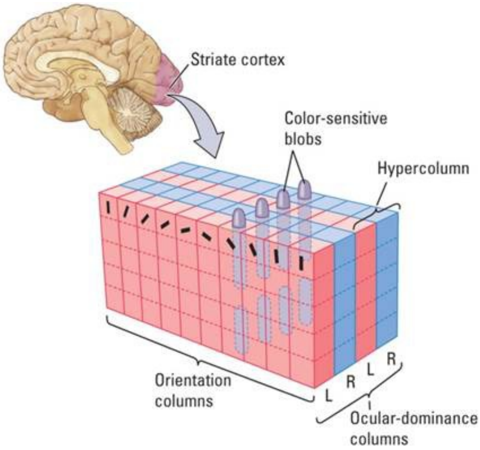
1. The human visual pathway



Left visual field
Right visual field
Nasal retina
Optical lens
Eye
Temporal retina
Temporal retina
Optic nerve
Optic chiasma
Lateral geniculate nucleus (LGN)
Pretectal nucleus
Primary visual cortex


A) Right visual hemifield
B) Left visual cortex
72°
50°
Upper
vertical
meridian
32°
24°
16°
8°
4°
Fovea
Lower
vertical
meridian
vertical
meridian
Lower
Fovea
MC
1°
2°
Upper
4°
vertical
8°
meridian
16°
24°
32°
50°
72°
5 mm
2.5 2.0 1.5 1 0.7 0.5
Anisotropy scale, (Mp/Me)

In 1962 Hubel and Wiesel propose the ice-cube model for V1 (Nobel Prize 1980):

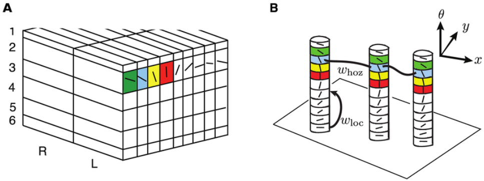


D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol, 160:106–154, 1962.

2. Mathematical modeling of the visual cortex

Visual cortex as a fiber (contact) bundle

In 1989 Hoffman (Caltech) proposes a model for V1 as contact bundle.

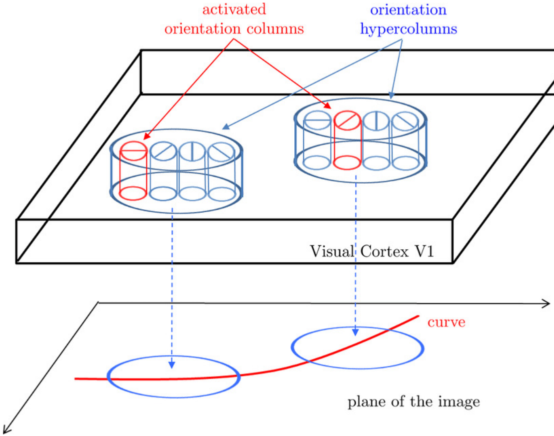


V1 as global S^1 fiber bundle on $\mathbb{R}^2$:

$$\mathbb{R}^2 \times S^1 \longrightarrow \mathbb{R}^2, \qquad (x, y), \theta \mapsto (x, y)$$

W.C. Hoffmann. The visual cortex is a contact bundle. mathematics and computation, 32:137–167, 1989.

How the visual cortex detects the contour of an image:



At each point of V1 we have the information regarding **all** possible directions: only the detected direction will be highlighted in the hypercolumn!

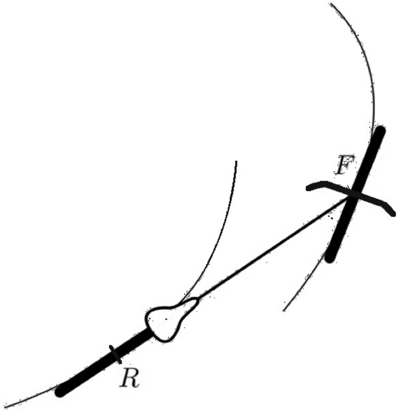
1. $\mathcal{R} : V \longrightarrow \mathbb{R}$ contour perceived on $V \subset \mathbb{R}^2$, base space of $V1 = \mathbb{R}^2 \times S^1$ from retina
2. **Hypercolumnar orientation detection**: scalar function

$$\begin{array}{cccc} \Theta: & V & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & \Theta(x,y) := \mathrm{argmax}_{\theta\in[0,2\pi]}\{X(\theta)\mathcal{R}(x,y)\} \end{array}$$

where $X = -\sin\theta\,\partial_x + \cos\theta\,\partial_y$ is a vector field on $V1 = S^1 \times \mathbb{R}^2$.

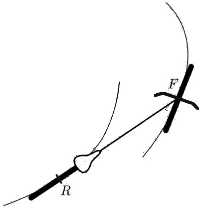
3. A sub Riemannian approach to the question of border completion

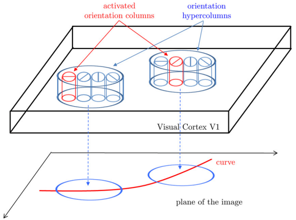
On a bicicle we have a constraint on the direction to take **not** the path.



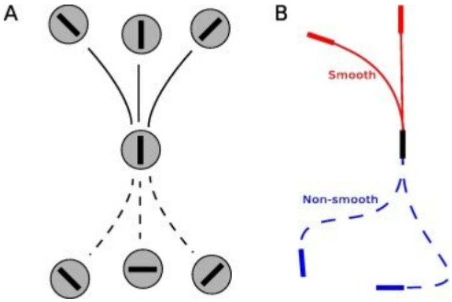
Dictionary

Mathematics	Bicycle	Primary visual cortex
$b = (x, y)$	position of the rear wheel	point in the image of the retinotopic map of the visual field
$v = f - b = (\cos\theta, \sin\theta)$	direction of motion	detected orientation at b
(b, v)	point in configuration space	hypercolumn
$Z = \dot{v}$	normal to rear wheel path	orientation vector field
$Q \cong \mathbb{R}^2 \times S^1 \cong \mathrm{SE}(2)$	configuration space of bicycle	V1 total space





How do we complete a non existing border?



David J. Field, Anthony Hayes, and Robert F. Hess. Contour integration by the human visual system: Evidence for a local “association field”. Vision Research, 33(2):173–193, 1993.

The problem of border reconstruction via Sub-Riemannian metric
Idea: we build a geodesic on the whole space according to some metric and then we **project** it on the distribution.

$$(x, y, \theta) \mapsto \mathcal{D}_{(x,y,\theta)} = \mathrm{span} \left\{ \begin{matrix} X_1 = \cos\theta\,\partial_x + \sin\theta\,\partial_y \\ X_2 = \partial_\theta \end{matrix} \right\}$$

We want to find curves $\gamma(t)$ that are **tangent** to the distribution:

$$\gamma'(t) \in \mathrm{span} \left\{ \begin{matrix} X_1 = \cos\theta\,\partial_x + \sin\theta\,\partial_y \\ X_2 = \partial_\theta \end{matrix} \right\}$$

They will be the geodesics in a subriemannian metric!
How to find them: Hamilton equations!

Sub-riemannian geodesics

Hamilton Equations for Subriemannian geodesics:

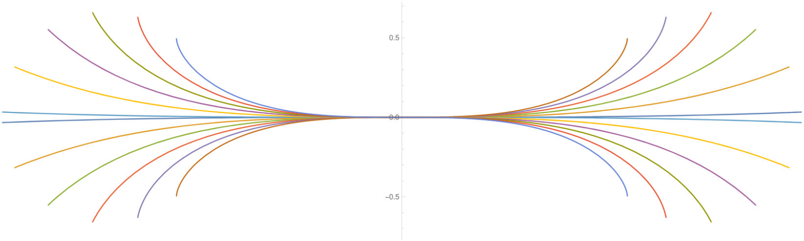
$$\begin{cases} \dot{x} = \cos\theta\, p_1 & \dot{p}_1 = p_3\, p_1 \\ \dot{y} = \sin\theta\, p_1 & \dot{p}_2 = -p_3\, p_1 \\ \dot{\theta} = p_2 & \dot{p}_3 = 0 \end{cases}$$

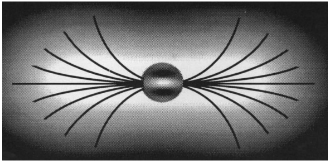
Geodesic solutions, with 6 parameters to be determined from the initial conditions

$$x(t) = \int_0^t v \cos(\omega s \phi) \cos(\theta(s))\, ds + x_0$$

$$y(t) = \pm \int_0^t v \cos(\omega s \phi) \sin(\theta(s))\, ds + y_0$$

$$\theta(t) = \mp \frac{v}{\omega} \cos(\omega s \phi) + \theta_0$$


0.5
0.0
−0.5

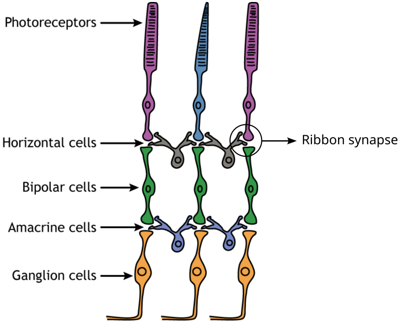


4. A Deep Learning approach to border enhancement and light sensitivity

The retina ganglions are capable of accounting for large light variations in images.

Barn owl (*Tyto alba*)

Idea: study the mathematical modeling for lower visual system to implement a Deep Learning neural network



The corresponding between retina and ganglionic layer is bijective:
Every point receives information about a *neighbourhood* of a point in R.

- **Bijection between Retina and Ganglionic layers:**

$$E \xrightarrow{G} E' \tag{1}$$

- **Receptorial and ganglionic activation functions**

$$\begin{aligned} \mathcal{R} : E &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto \mathcal{R}(x,y) \end{aligned} \qquad \begin{aligned} \mathcal{R}' : E' &\longrightarrow \mathbb{R} \\ (x',y') &\longmapsto \mathcal{R}'(x',y') \end{aligned}$$

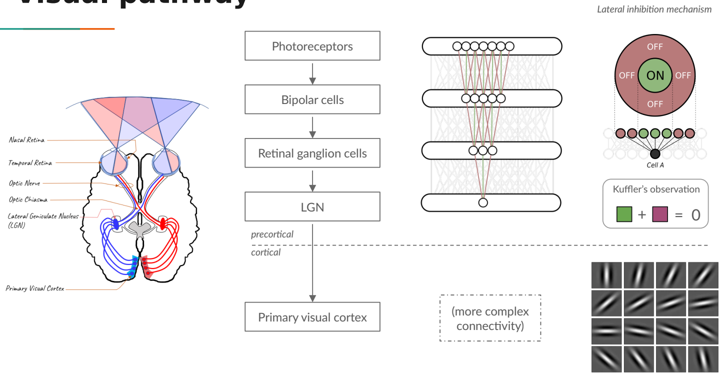
where we model the ganglionic activation as

$$\mathcal{R}'(x',y') = \int_{U_\rho(x,y)} \mathcal{R}(u,v)\, du\, dv$$

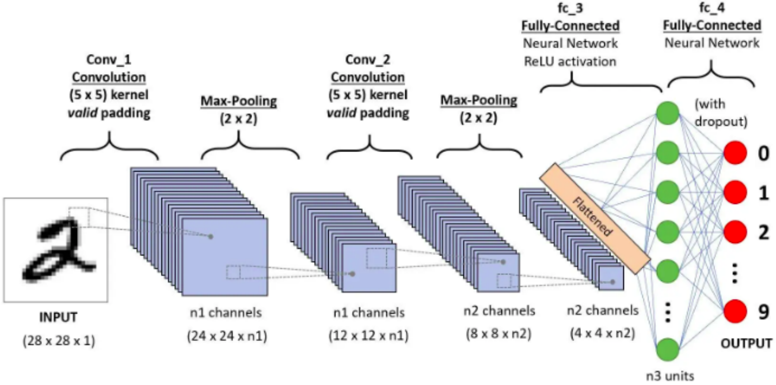
with $G(x,y) = (x',y')$ and

$$U_\rho(x,y) = \{(u,v) \in \mathbb{R}^2 : (u-x)^2 + (v-y)^2 \leq \rho^2\}$$

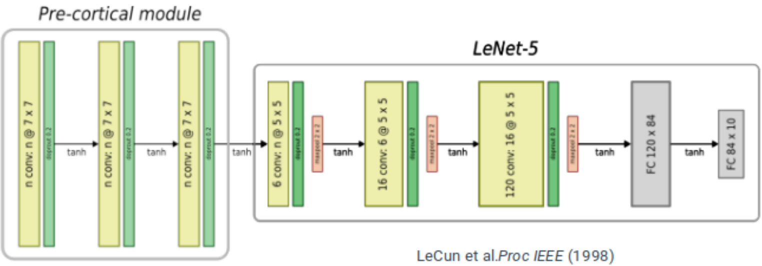
Visual pathway



Convolutional Neural Network Models



Implementation



(147 parameters)

LeCun et al.*Proc IEEE* (1998)

(~ 60 000 parameters)

The dataset used to test our net (Retilenet)

Testing outside of training "illumination statistics"

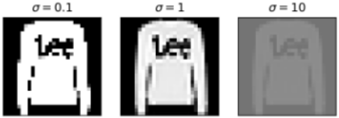
Consider an image $\vec{x}$, with pixels x_i. We define the **brightness change** as the mean offset

$$x_i \to x_i - \mu$$

and the **contrast change** as the rescaling

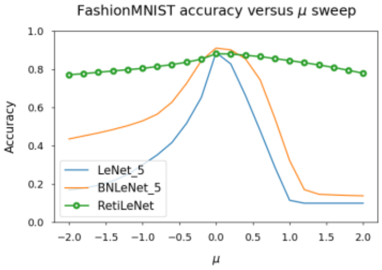
$$x_i \to \frac{x_i - \langle \vec{x} \rangle}{\sigma} + \langle \vec{x} \rangle$$

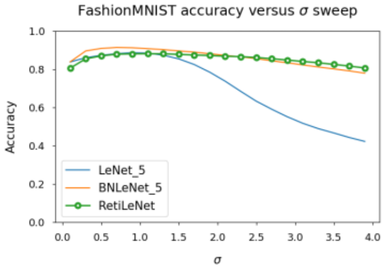
Brightness change



Contrast change

Results





Dataset: FashionMNIST

Model	μ -2.0	μ 0	μ 2.0	σ 0.1	σ 1.0	σ 3.9
LeNet-5	0.168	0.887	0.100	0.836	0.887	0.422
BNLeNet-5	0.435	**0.910**	0.139	**0.838**	**0.907**	0.779
RetiLeNet	**0.770**	0.880	**0.781**	0.805	0.880	**0.806**

Conclusions

We provide a sound mathematical modeling for the visual path for two purposes:

- Find a solution via Hamiltonian equations of the border completion problem.
- Implement a “precortical module” in a Deep Learning algorithm.

  1. We make a network invariant for a specific transformation by altering its structure
  2. This invariance can emerge spontaneously, without altering the training
  3. The modification of the net consists of a minimal number of weights