Semi-supervised Neural Networks solve an inverse problem for modeling Covid-19 spread

Alessandro Paticchio†, Tommaso Scarlatti†, Marios Mattheakis‡, Pavlos Protopapas‡ and Marco Brambilla†

†Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano

‡John A. Paulson School of Engineering and Applied Sciences, Harvard University

Abstract

Studying the dynamics of COVID-19 is of paramount importance to understanding the efficiency of restrictive measures and develop strategies to defend against up-coming contagion waves. We study the spread of COVID-19 using a semi-supervised neural network and assuming a passive part of the population remains isolated from the virus dynamics. We start with an unsupervised neural network that learns solutions of differential equations for different modeling parameters and initial conditions. A supervised method then solves the inverse problem by estimating the optimal conditions that generate functions to fit the data for those infected by, recovered from, and deceased due to COVID-19. This semi-supervised approach incorporates real data to determine the evolution of the spread, the passive population, and the basic reproduction number for different countries.

Methodology

- **Unsupervised part:** data-free NN trained to discover solutions for a DE system in a high-dimensional parametric space made of modeling-parameters and initial conditions [4].
- **Supervised part:** employs a gradient descent optimization method to determine the model parameters and initial conditions that best describe ground truth observations using automatic differentiation.
- **Parametrization:** Impose the initial conditions $\mathbf{z}_0$ in the predictions $\hat{\mathbf{z}}$. [5].

$$\frac{d\mathbf{z}}{dt} = g(\mathbf{z}), \quad \text{with} \quad \mathbf{z}(t=0) = \mathbf{z}_0, \tag{1}$$

$$\hat{\mathbf{z}} = \mathbf{z}_0 + f(t)\,(\mathbf{z}_{\mathrm{NN}} - \mathbf{z}_0), \quad \text{where} \quad f(t) = 1 - e^{-t} \tag{2}$$

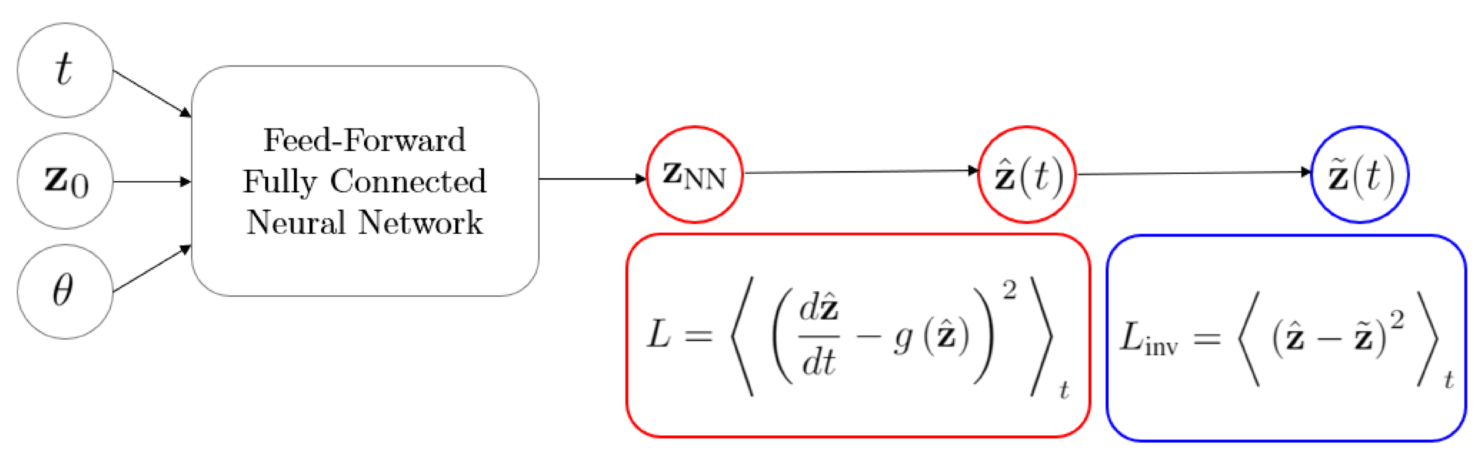

Fig. 1: Semi-supervised neural network architecture. Red and blue indicate, respectively, the unsupervised and supervised learning parts.

- **SIR model:** Assess the performance of the proposed method by studying the simple SIR model with a loss function [2, 3]

$$L = \left\langle \left(\frac{d\hat{S}}{dt} + \beta\hat{S}\hat{I}\right)^2 + \left(\frac{d\hat{I}}{dt} - \beta\hat{S}\hat{I} + \gamma\hat{I}\right)^2 + \left(\frac{d\hat{R}}{dt} - \gamma\hat{I}\right)^2 \right\rangle_t. \tag{3}$$

SIR Analysis

- **Network architecture**: 4 hidden layers with 50 neurons per hidden layer. Trained for for $2 \cdot 10^4$ epochs with Adam optimizer with learning rate $8 \cdot 10^{-4}$.
- **Network optimization**: Training loss function for two different architectures

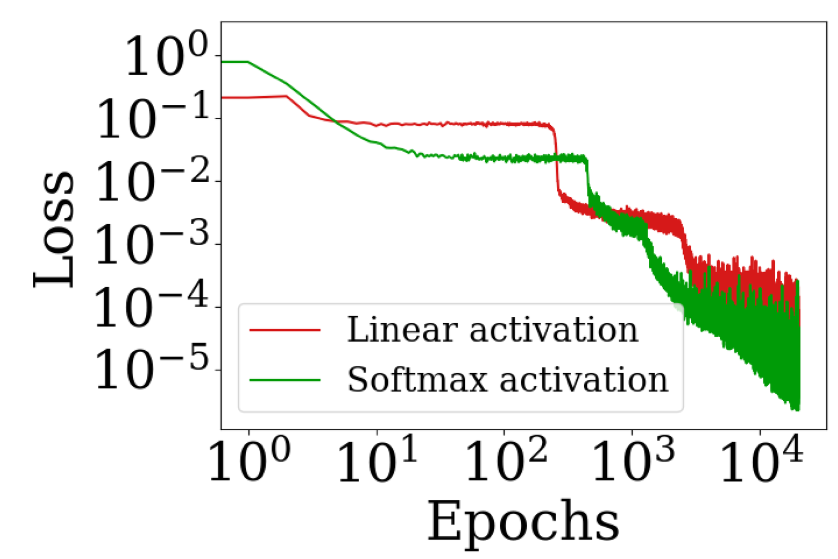

Fig. 2: Training with softmax (green) and identity (red) activation functions in the last layer. The oscillations in the curve are due to the perturbation applied to the training points, and it is visually amplified by logarithmic scale of the y-axis.

- **Error analysis:** Smaller validation loss reached inside the training bundle (dashed box) and gradually increases moving out.

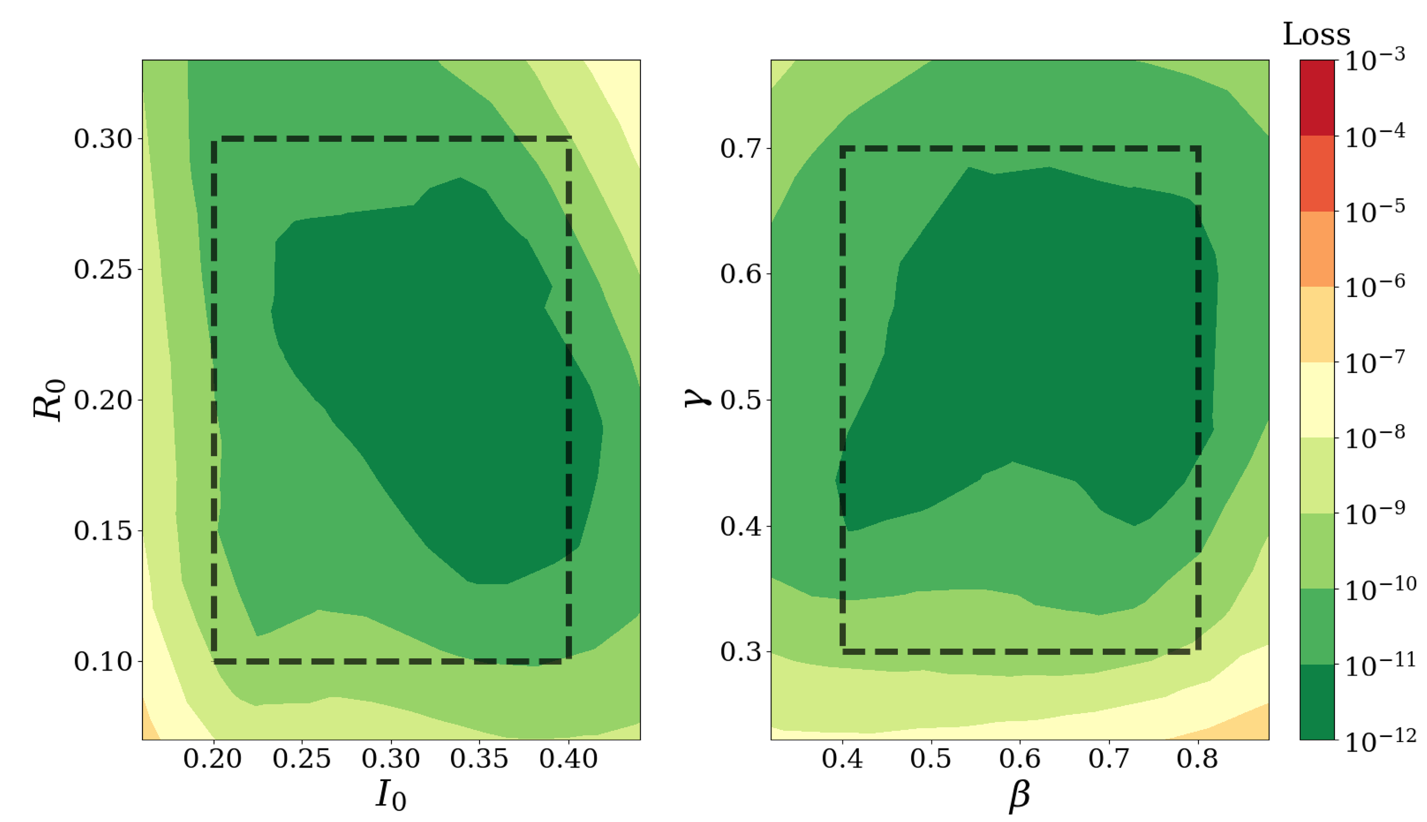

Fig. 3: Validation loss of the model in different areas, as a function (a) of the initial conditions, and (b) of the parameters. The area in the black dashed-line boxes is the training bundles regime.

References

[1] Jonas Dehning et al. In: *Science* (2020).
[2] Klaus Dietz. In: *Statistical methods in medical research* (1993).
[3] Pauline [van den Driessche]. In: *Infectious Disease Modelling* (2017).
[4] Cedric Flamant, Pavlos Protopapas, and David Sondak. 2020. arXiv: `2006.14372 [physics.comp-ph]`.
[5] Marios Mattheakis et al. In: *arXiv* (2020). arXiv: `2001.11107 [physics.comp-ph]`.

COVID-19: Real data

- **SIRP**: Consider a passive compartment P isolated from the dynamics.

$$N = S + I + R + P \tag{4}$$

$$L = \left\langle \left(\frac{d\hat{S}}{dt} + \beta\hat{S}\hat{I}\right)^2 + \left(\frac{d\hat{I}}{dt} - \beta\hat{S}\hat{I} + \gamma\hat{I}\right)^2 + \left(\frac{d\hat{R}}{dt} - \gamma\hat{I}\right)^2 + \left(\frac{d\hat{P}}{dt}\right)^2 \right\rangle_t \tag{5}$$

- Train the model in the lockdown period for each country, assuming constant modeling parameters [1].

Country	I_0	R_0	P_0	β	γ
Switzerland	[0.01, 0.02]	[0.001, 0.006]	[0.9, 0.97]	[0.7, 0.9]	[0.15, 0.3]
Spain	[0.01, 0.02]	[0.004, 0.009]	[0.9, 0.97]	[0.4, 0.6]	[0.1, 0.2]
Italy	[0.01, 0.02]	[0.001, 0.006]	[0.9, 0.97]	[0.4, 0.6]	[0.1, 0.2]

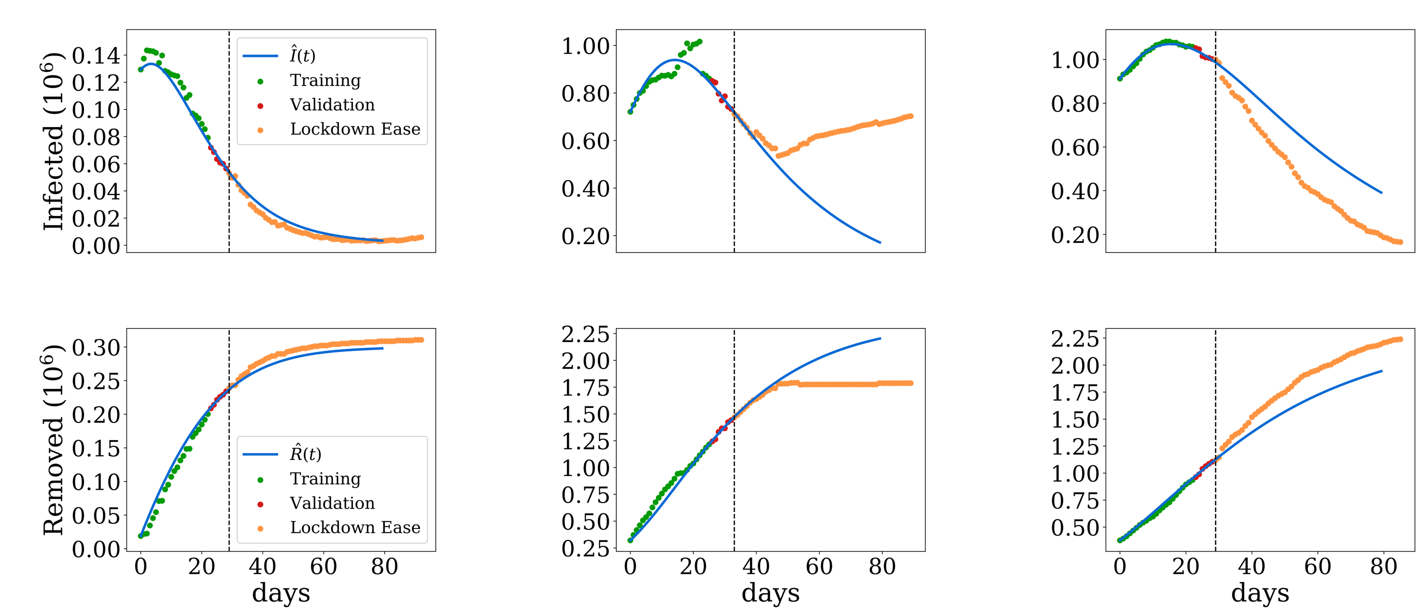

Fig. 4: Infected and removed populations for Switzerland (left column), Spain (middle), Italy (right). Points indicate data, solid lines denote predictions, dashed lines show the end of lockdown.

- Italy has been the most impacted country, reaching a reproduction number $\mathfrak{R}_0 = 4.7$ with $P = 96\%$. Spain follows with $\mathfrak{R}_0 = 3.3$ and $P = 95\%$. Switzerland has also $P = 96\%$, with the smallest $\mathfrak{R}_0$, resulting in $\mathfrak{R}_0 = 2.7$.

Conclusions

- Epidemiology-informed network incorporates real data to study the COVID-19 spread in Switzerland, Spain, and Italy.
- Semi-supervised neural networks solve inverse problems formulated by DEs.
- Learning a family of solutions for nonlinear systems.
- Solving the inverse problem determines modeling parameters yielding predicted solutions that best fit the data.
- SIRP model captures the dynamics of COVID-19 spread.