Human or Machine?
It Is Not What You Write, But How You Write It

Luis A. Leiva
Aalto University, Finland
name.surname@aalto.fi

Moises Diaz
Universidad del
Atlantico Medio, Spain
moises.diaz@atlanticomedio.es

Miguel A. Ferrer
Universidad de Las Palmas
de Gran Canaria, Spain
miguelangel.ferrer@ulpgc.es

Réjean Plamondon
Polytechnique Montréal, Canada
rejean.plamondon@polymtl.ca

Introduction

The **liveness detection** problem: *How to ensure a human is operating a computer application?*

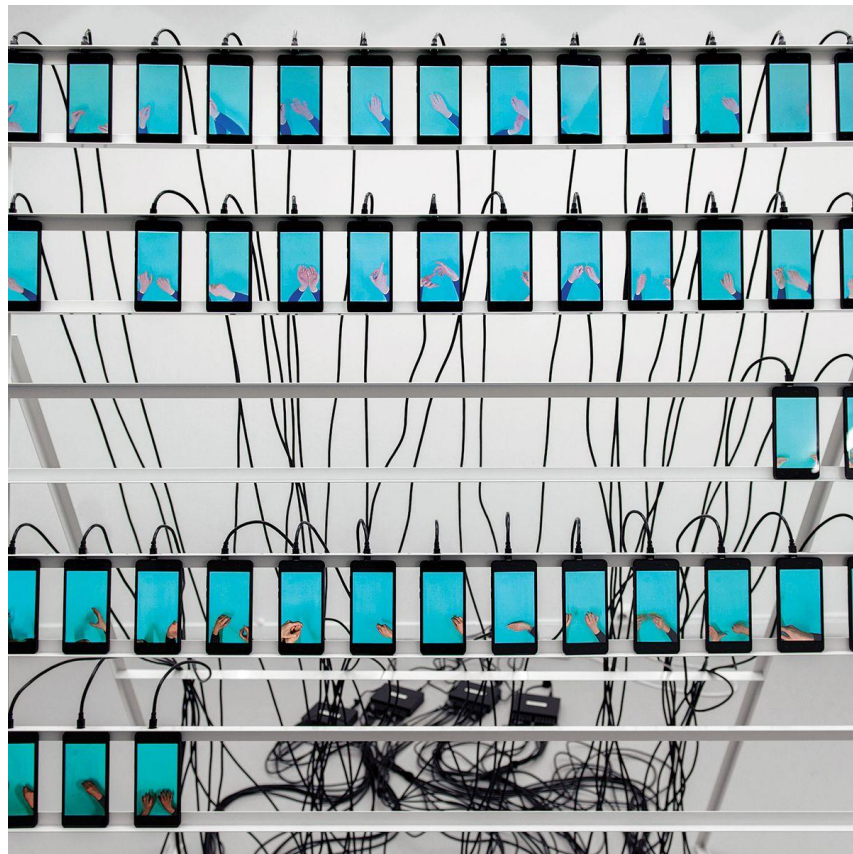We classify handwriting movements as human or machine-generated using Deep Learning models.
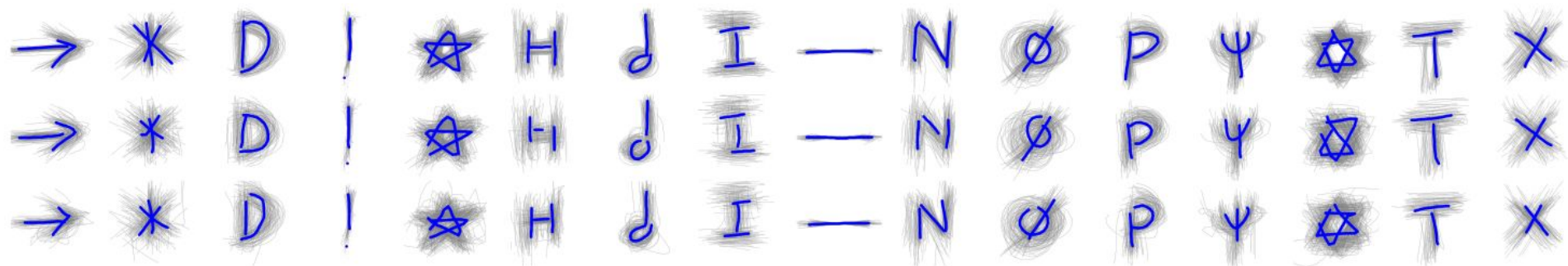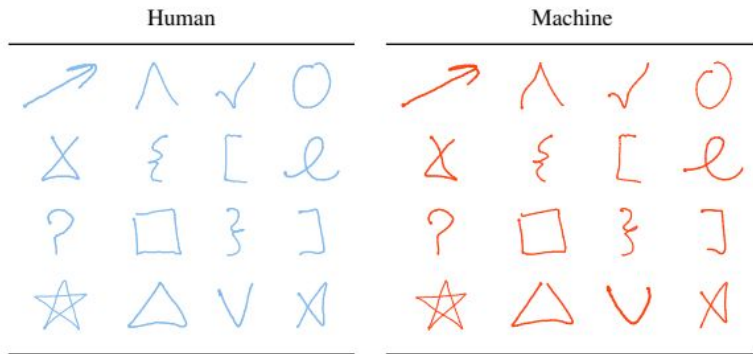

Image credit: https://nymag.com/

Previous work

Human likeness experiments with **signatures** [Galbally et al. 2012] and **gestures** [Leiva et al. 2016, 2017]
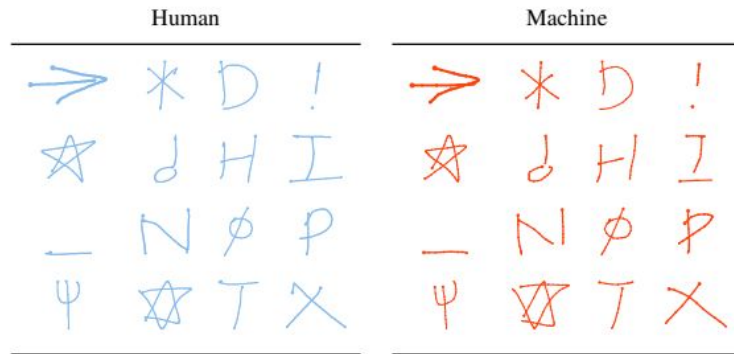
→ They showed images to the users, thereby ignoring movement dynamics.

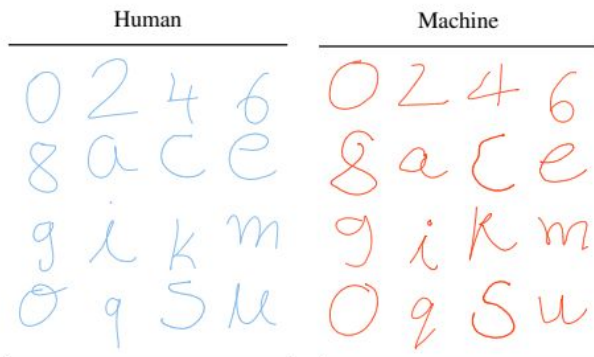→ We analyze movements both as images and spatiotemporal trajectories.
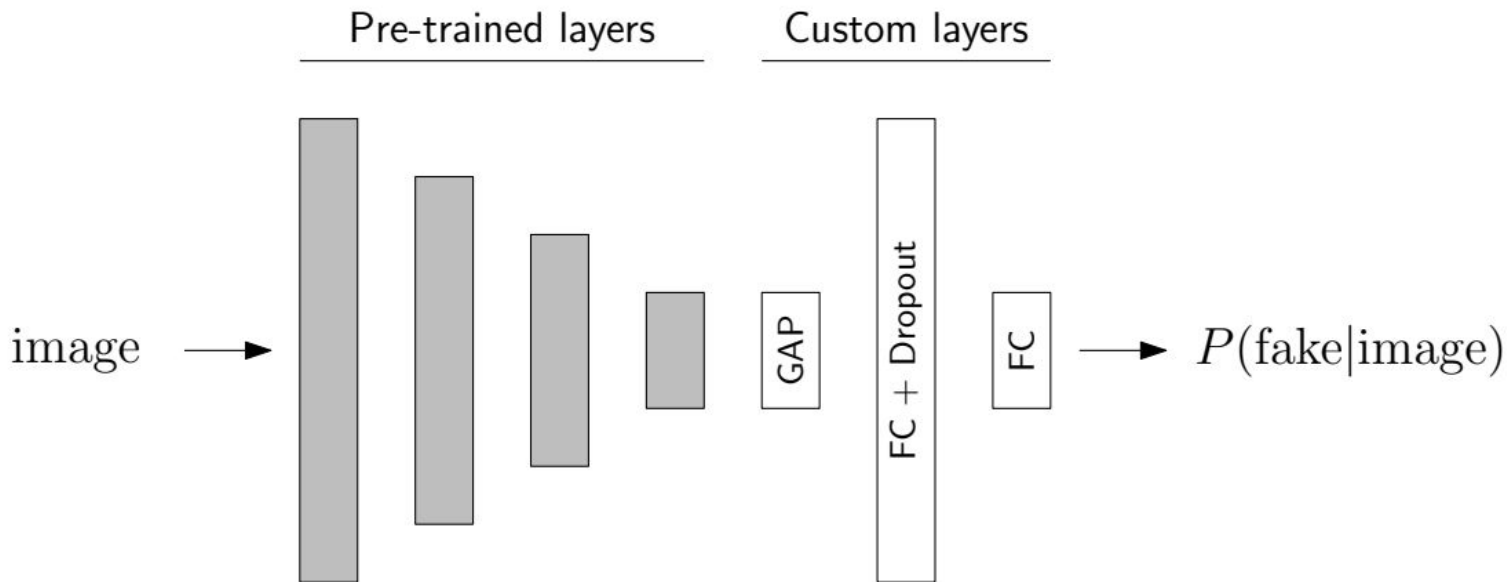
Evaluated datasets

(a) $1-GDS dataset examples

(b) $N-MMG dataset examples

(c) Chars74k dataset examples
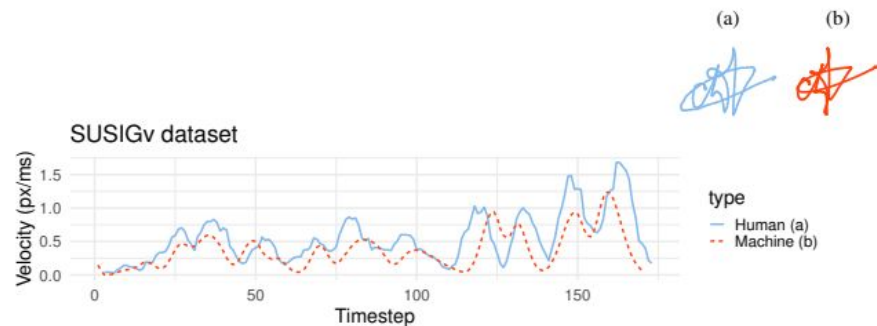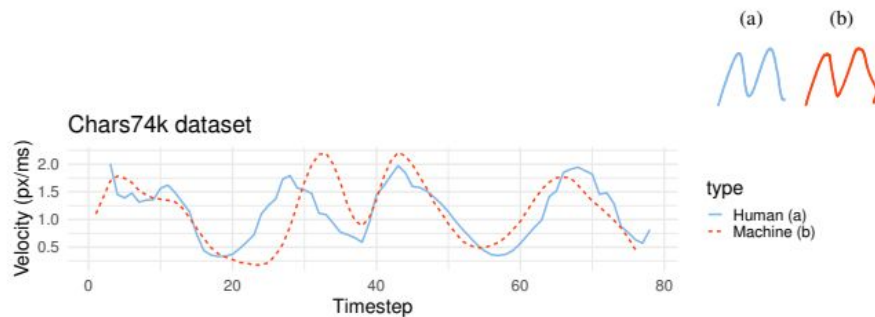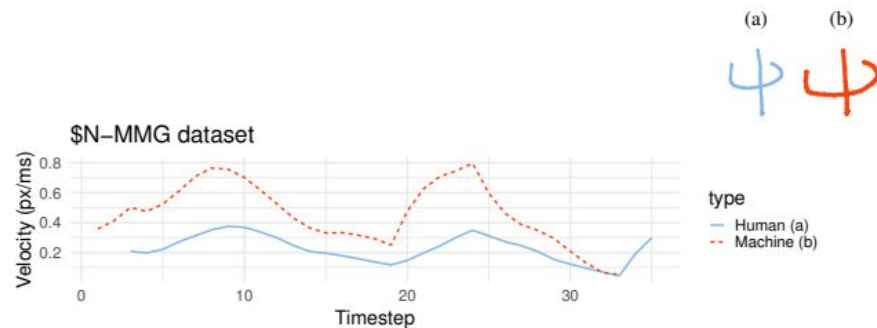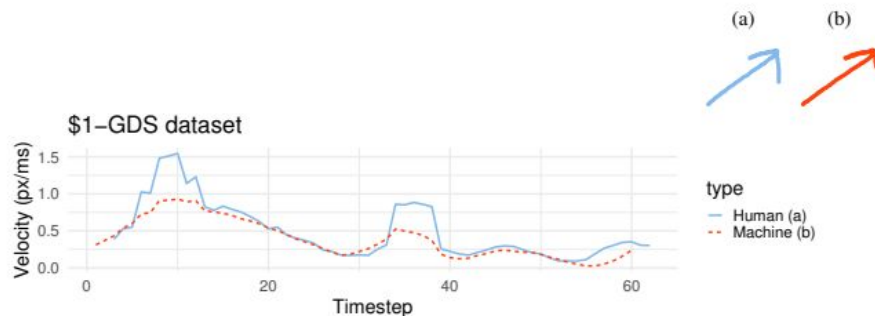
(d) SUSIGv dataset examples
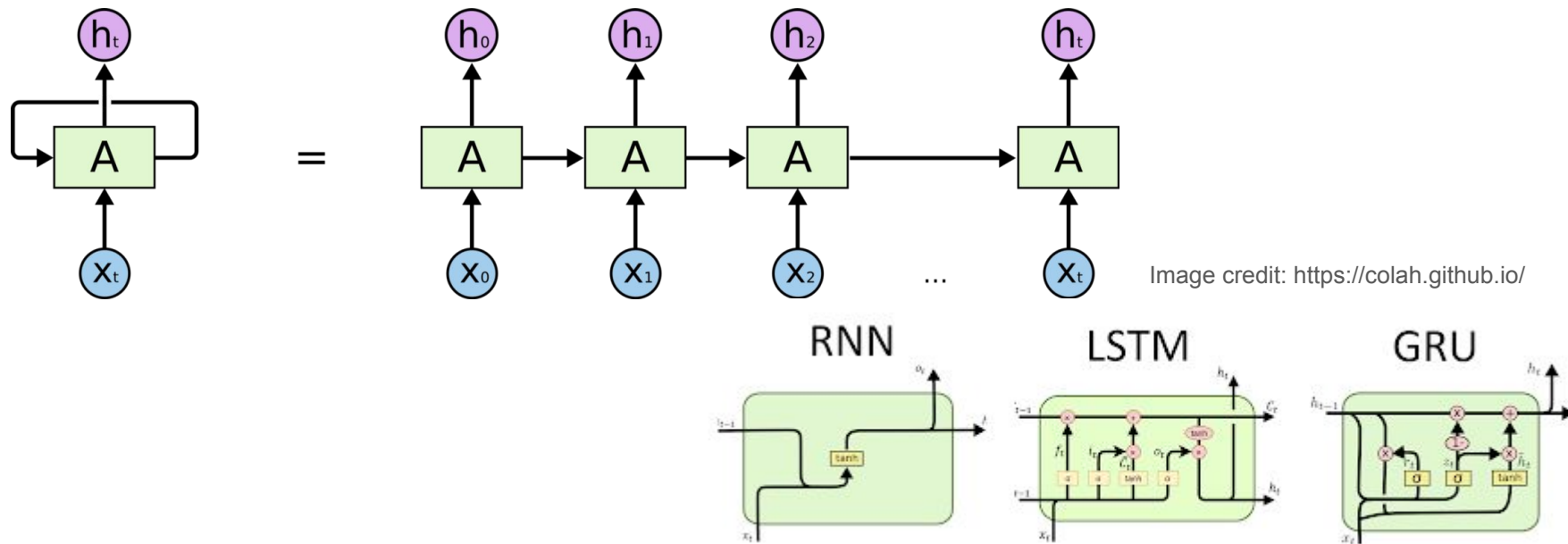
Convolutional Neural Nets (CNNs)

We do **transfer learning** with the following pre-trained models: VGG16, ResNet50, DenseNet, Inception, Xception

Velocity profile examples

Recurrent Neural Nets (RNNs)

Image credit: https://colah.github.io/

We **train from scratch** Vanilla RNN, LSTM, and GRU models (including their bi-directional variants)

Results with the $1-GDS dataset

System	Precision	Recall	F-measure	Accuracy	AUC
VGG16	68.50	68.48	68.45	68.48	75.39
ResNet50	68.97	68.95	68.95	68.95	76.34
Xception	71.15	71.03	70.98	71.03	79.58
DenseNet	71.41	71.41	71.41	71.41	78.74
Inception	**75.09**	**74.69**	**74.57**	**74.69**	**82.82**
Custom CNN	74.32	74.28	74.27	74.28	82.50
1NN-DTW	85.32	83.97	83.80	83.97	83.88
Vanilla RNN	94.93	94.51	94.49	94.51	94.46
LSTM	95.66	95.27	95.25	95.27	97.45
Bi-LSTM	95.14	94.73	94.72	94.73	96.98
GRU	**95.78**	**95.39**	**95.38**	**95.39**	**98.20**
Bi-GRU	95.62	95.20	95.19	95.20	97.76

Results with the GRU model (all datasets)

Dataset	Precision	Recall	F-measure	Accuracy	AUC
$1-GDS	95.78	95.39	95.38	95.39	98.20
$N-MMG	87.41	86.98	86.94	86.98	92.07
Chars74k	97.06	97.04	97.04	97.04	99.30
SUSIGv	93.68	93.35	93.34	93.35	95.43

See Table IV in our paper for additional results with the other RNNs.

Additional experiments available

- Deep learning models complexity
- RNN models performance on all datasets
- GRU robustness on all datasets
- Effect of input device

Please read our paper for more details!

Takeaway message

Movement dynamics reveal critical aspects "human likeness"

      → which are hard to simulate!

What really matters is not *what* you write but *how* you write it.