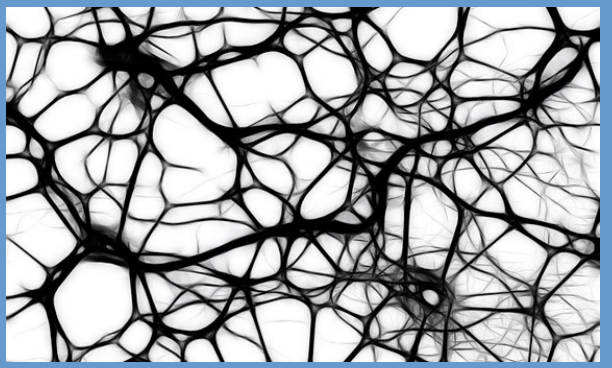
Deep Learning



From

Script and Address Recognition

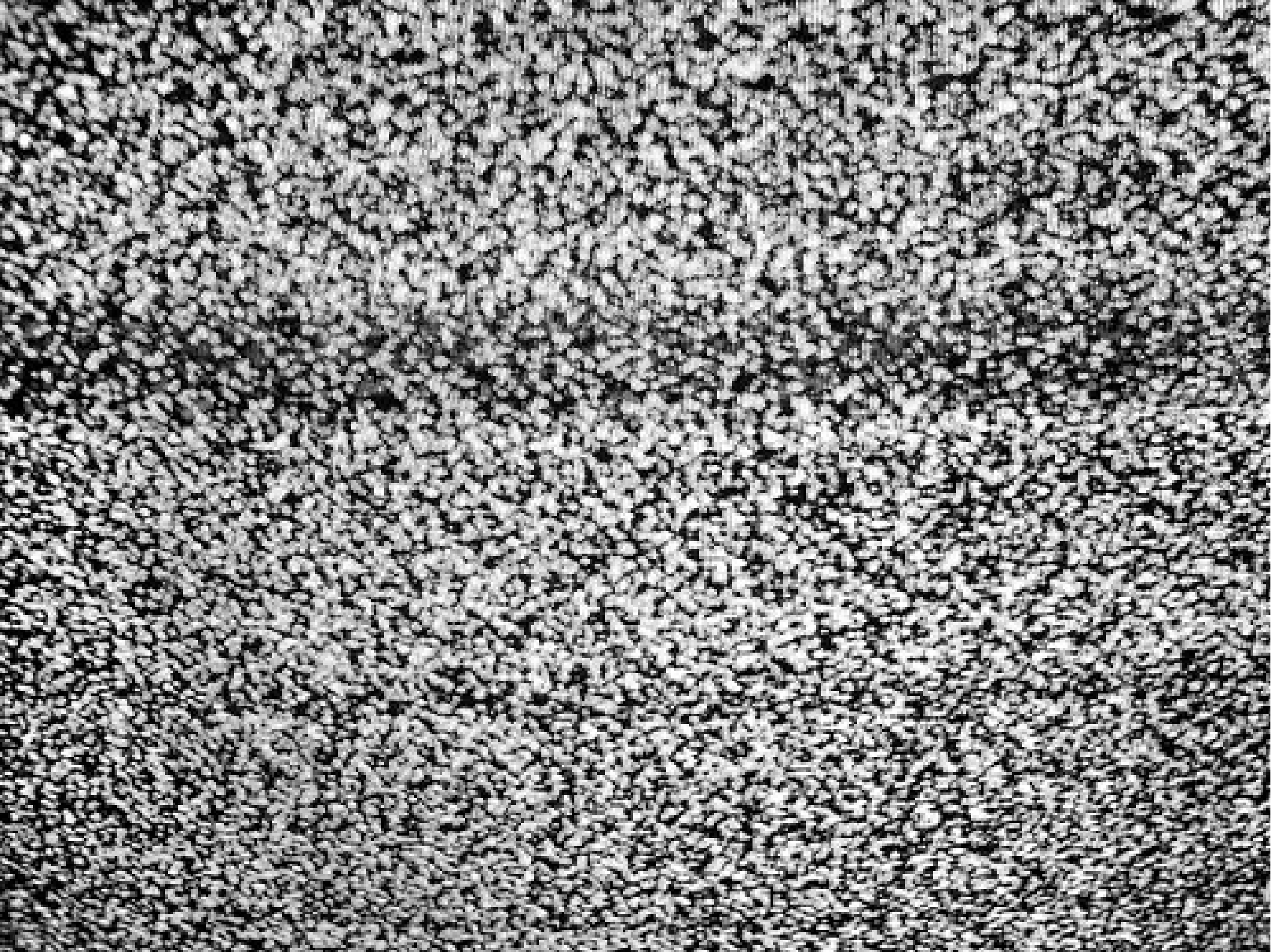
to

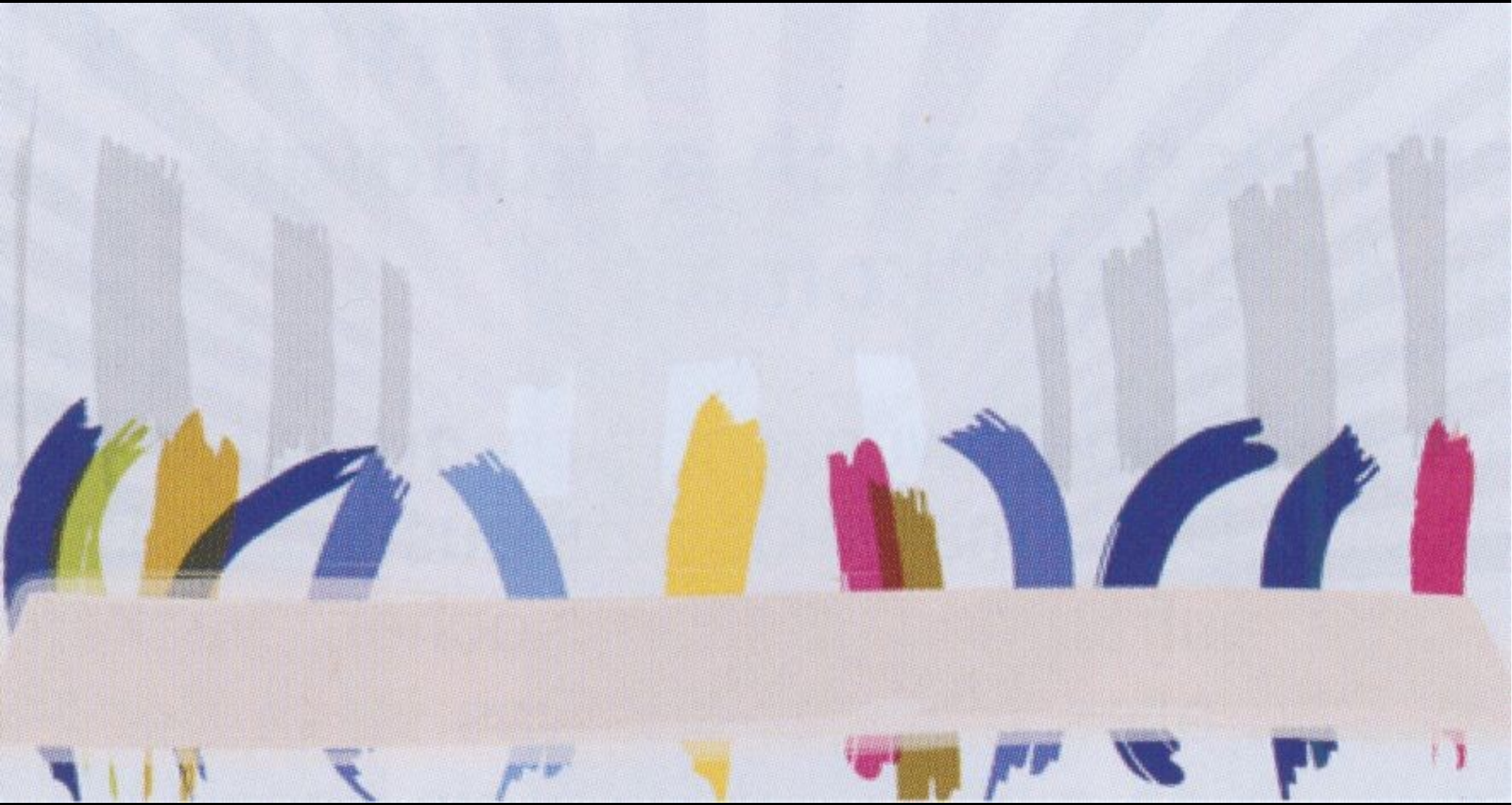
General Decisions and Solutions

Marc-Peter Schambach
*1971

Dipl. Physics
PhD Computer Science

Siemens Konstanz, Germany
Postal, Parcel & Airport Logistics
Handwriting Recognition





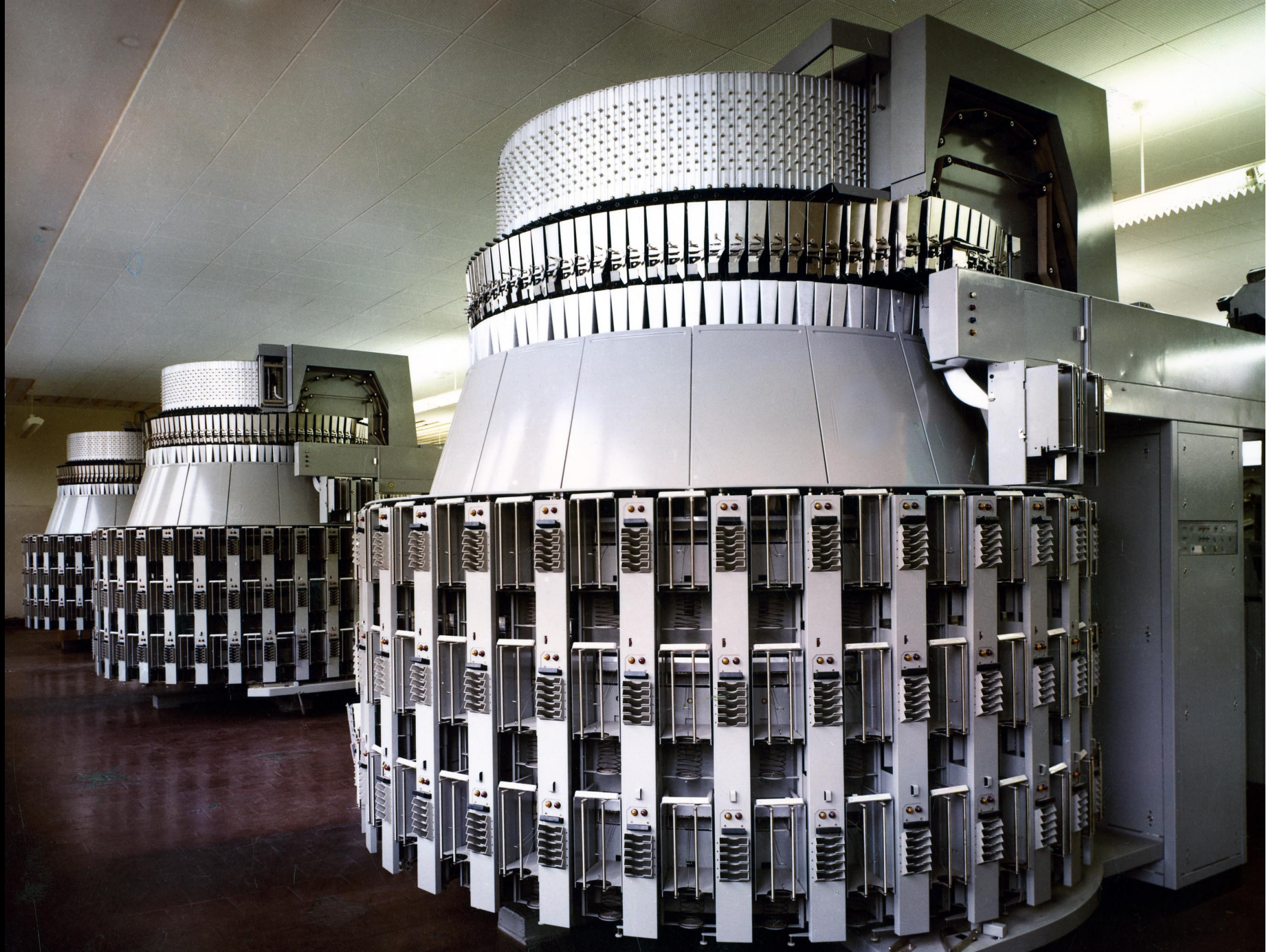
Overview

- Postal Automation
- Retinal Stimulus Reconstruction
- Handwriting Recognition
  - HMM
  - RNN
- Deep Learning
  - Geolocation of Images
  - Game of Go
  - Playing Atari

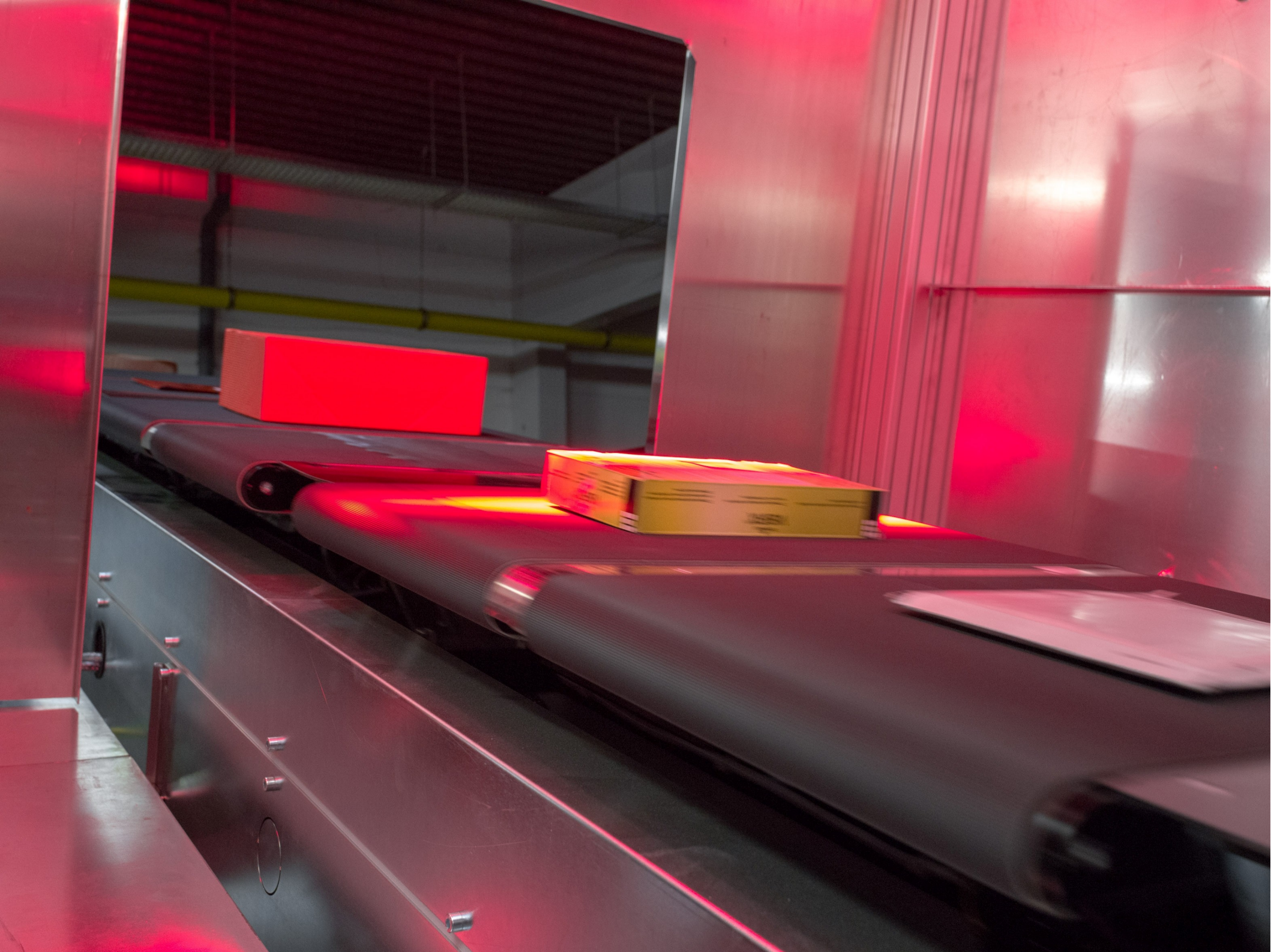
I

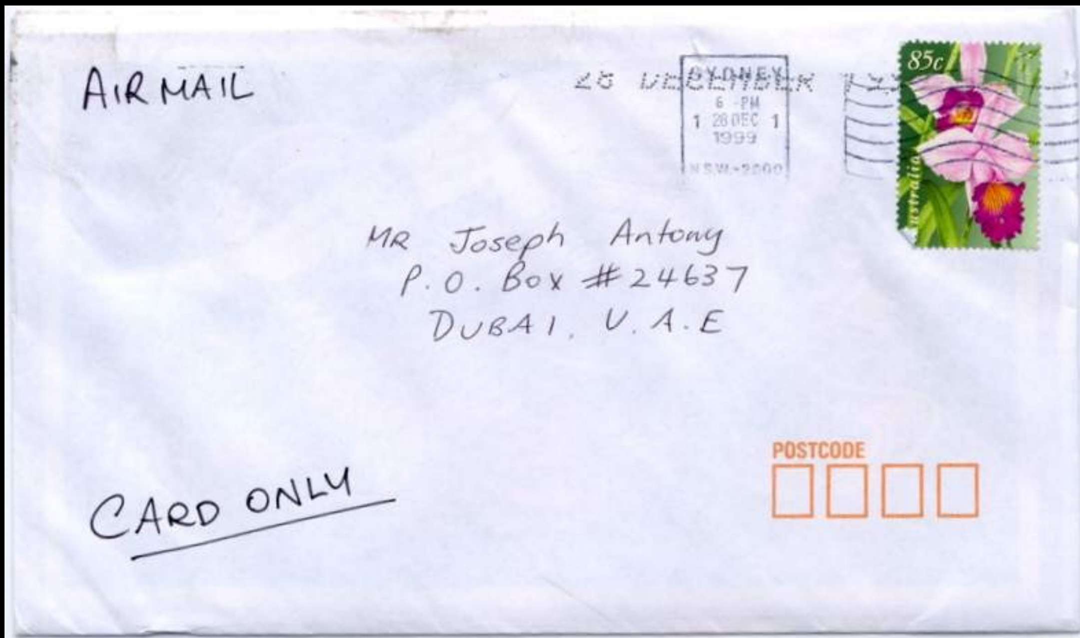
Postal Automation

Siemens Postal, Parcel & Airport Logistics GmbH

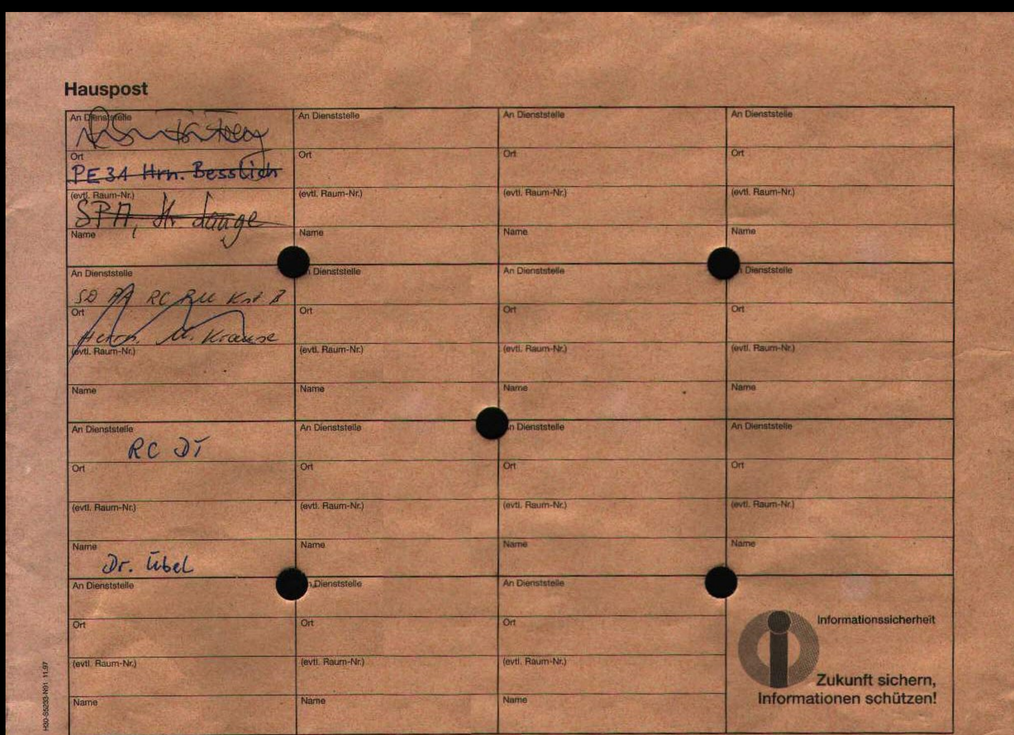



=ZP00BB01.RF501

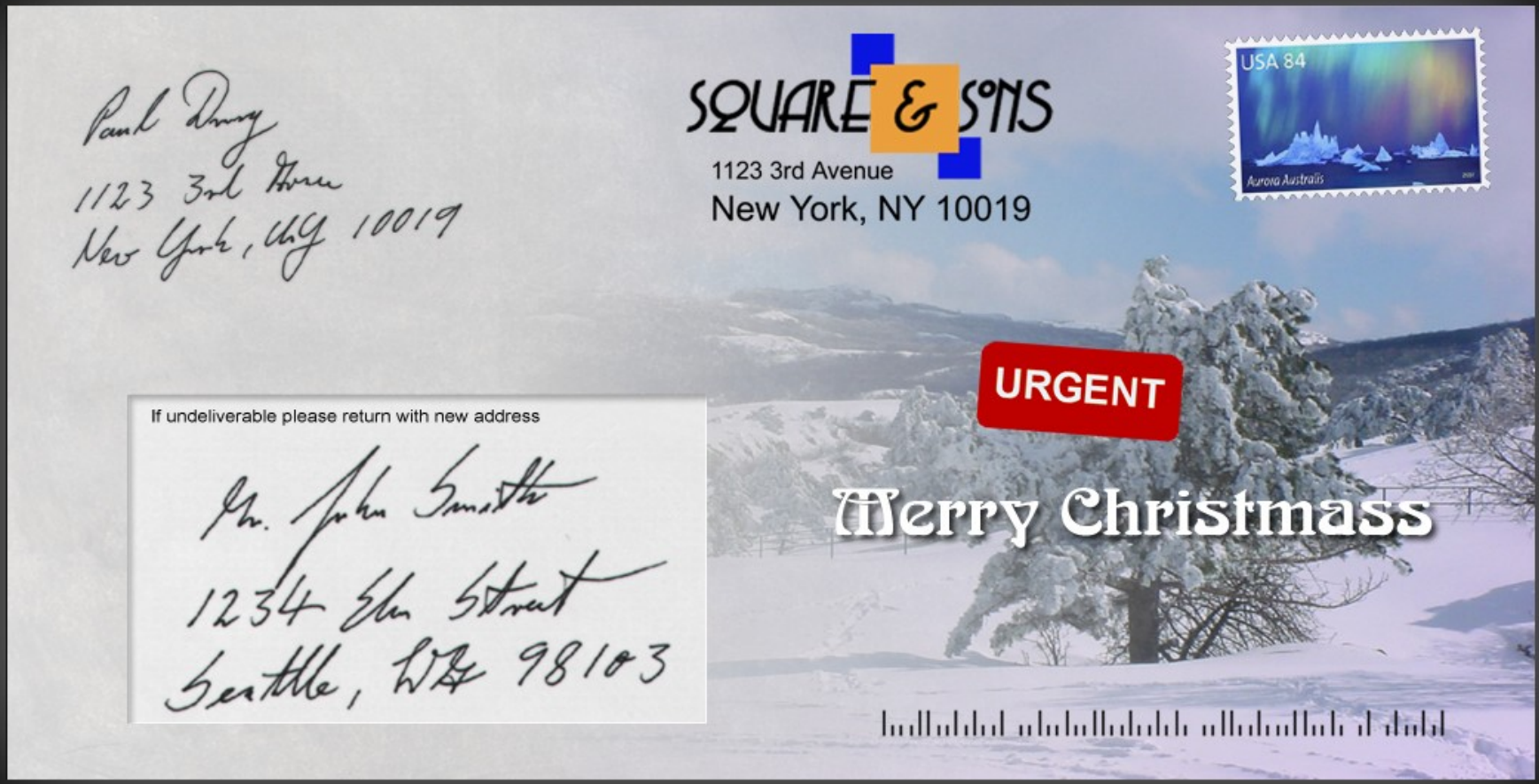



AIR MAIL
SYDNEY
6 -PM
1 28 DEC 1
1999
NSW-2000
85c
MR Joseph Antony
P.O. Box #24637
DUBAI, U.A.E
POSTCODE
CARD ONLY


HARD DRIVES
Get a Power User 540MB hard drive
only $349
CATHERINE G KILKENNY
OR CURRENT RESIDENT
1117 GLENN AVE
SAN JOSE CA 95125-3225
PAID
200MB Removable SyQuest Drive
Power Macintosh 6100/66
Special offer FileMaker Pro
QuickCam
The Visioneer PaperPort
Power Macintosh 9500
LAST CHANCE
Upgrade to Adobe Illustrator 5.5
NOW Up-to-Date & Contact 3.5
The Ultimate Desktop Publishing Bundle!
Doom II


Hauspost
An Dienststelle
Ort
(evtl. Raum-Nr.)
Name
PE 34 Hrn. Besstich
SPA, Hr. Lange
RC DT
Dr. Übel
Informationssicherheit
Zukunft sichern,
Informationen schützen!


99.34.137967.00000032
SI
RMP
Signature
Eigenhändig
A remettre en main propre
Da consegnare in mani proprie
Die Schweizerische Post
Viktoriastrasse 21
3030 Bern
Signor
Franco Rocco
Via Stazione 45
6500 Bellinzona
LA POSTE
DIE POST
LA POSTA


Paul Drury
1123 3rd Ave
New York, NY 10019
SQUARE & SONS
1123 3rd Avenue
New York, NY 10019
USA 84
Aurora Australis
URGENT
If undeliverable please return with new address
Mr. John Smith
1234 Elm Street
Seattle, WA 98103
Merry Christmass

Paul Drury
1123 3rd Ave
New York, NY 10019

SQUARE & SONS
1123 3rd Avenue
New York, NY 10019

URGENT

If undeliverable please return with new address

Mr. John Smith
1234 Elm Street
Seattle, WA 98103

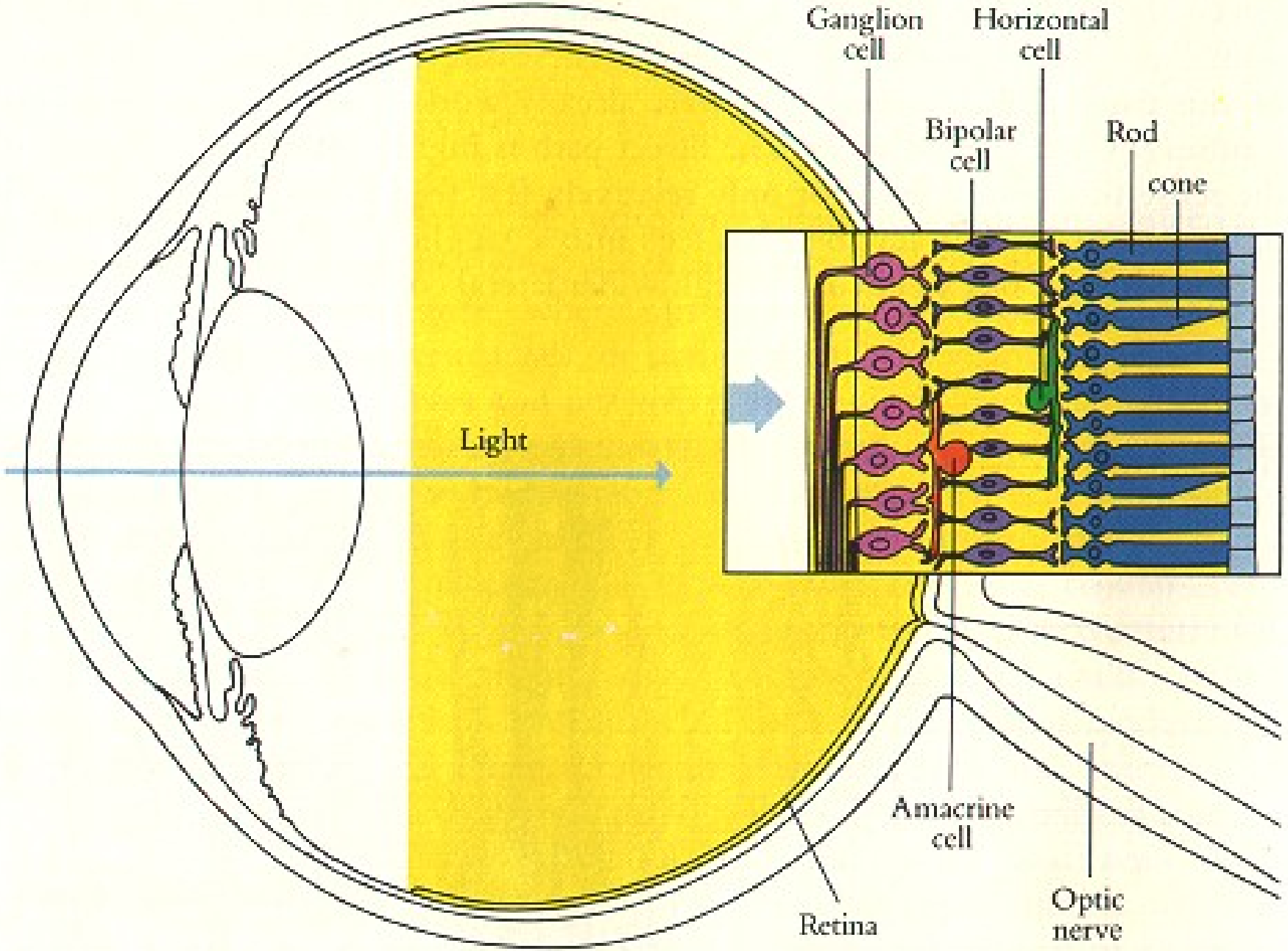
Merry Christmass

II

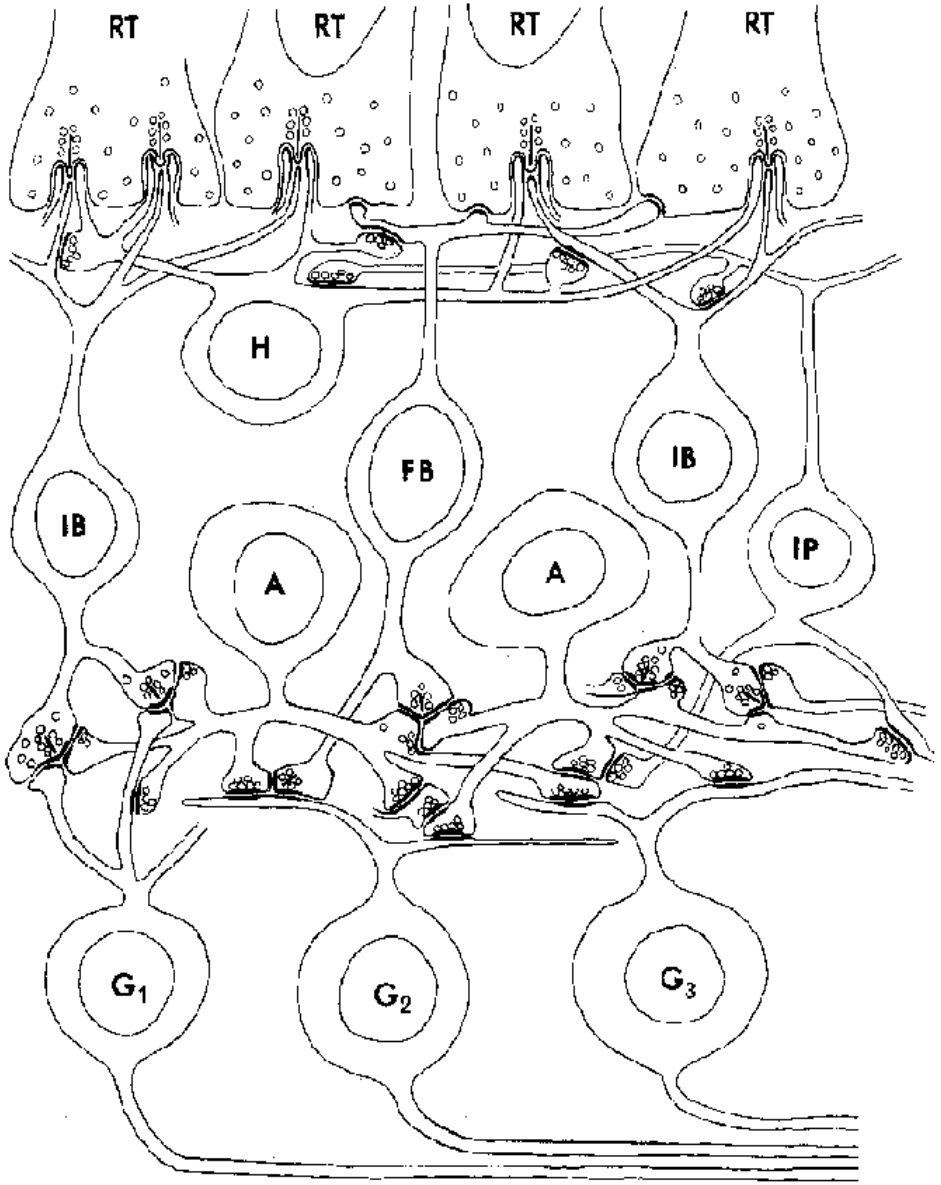
Neurobiology

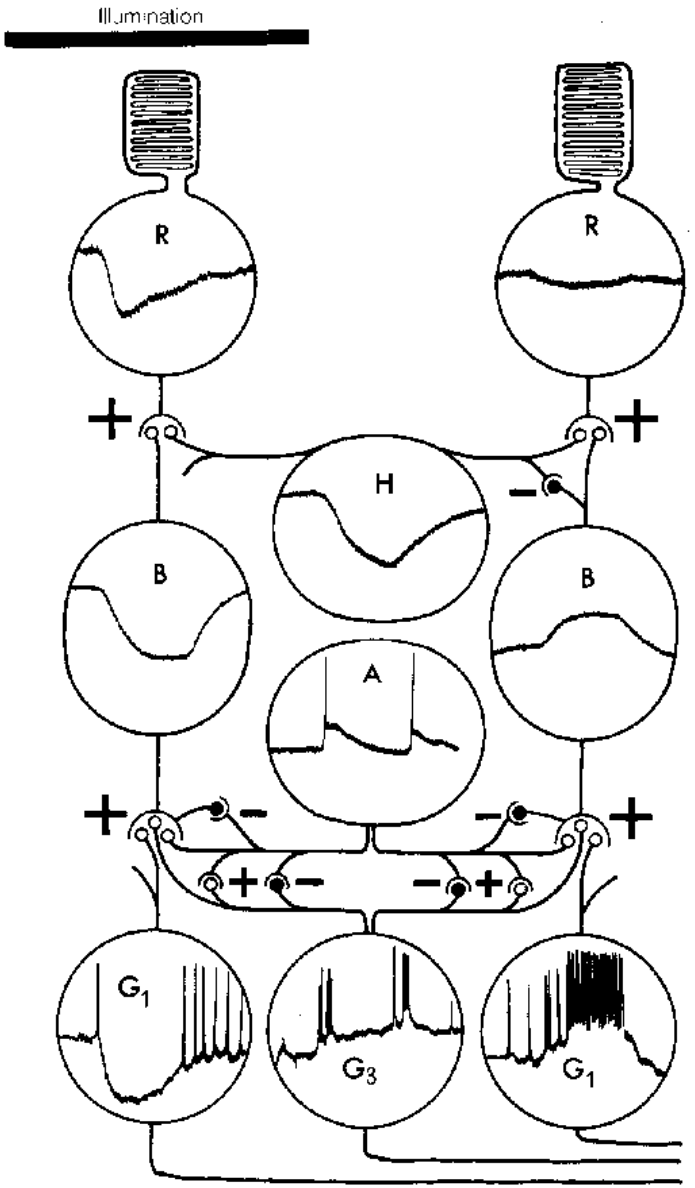
Retinal Stimulus Reconstruction

Retinal Stimulus Reconstruction

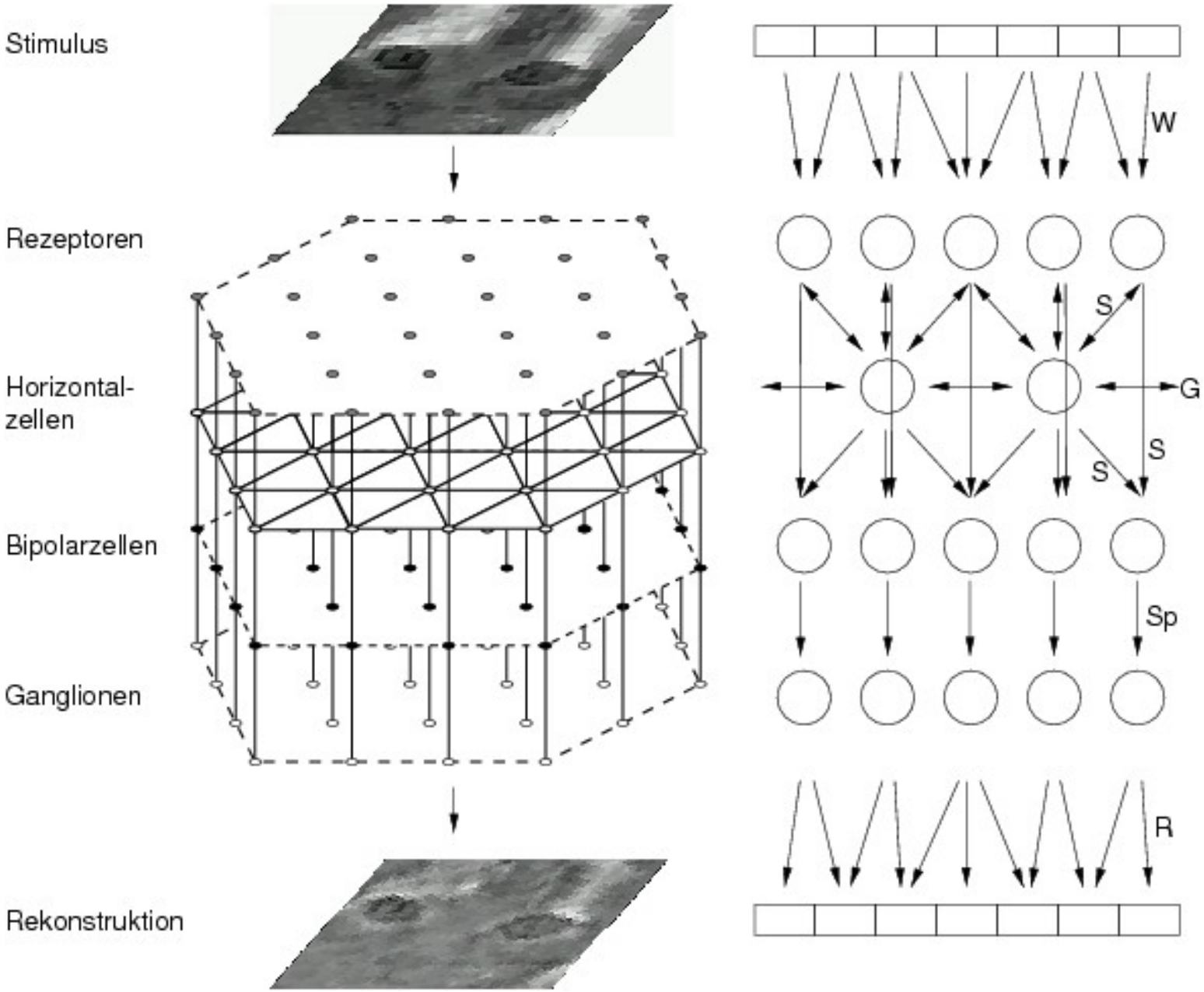


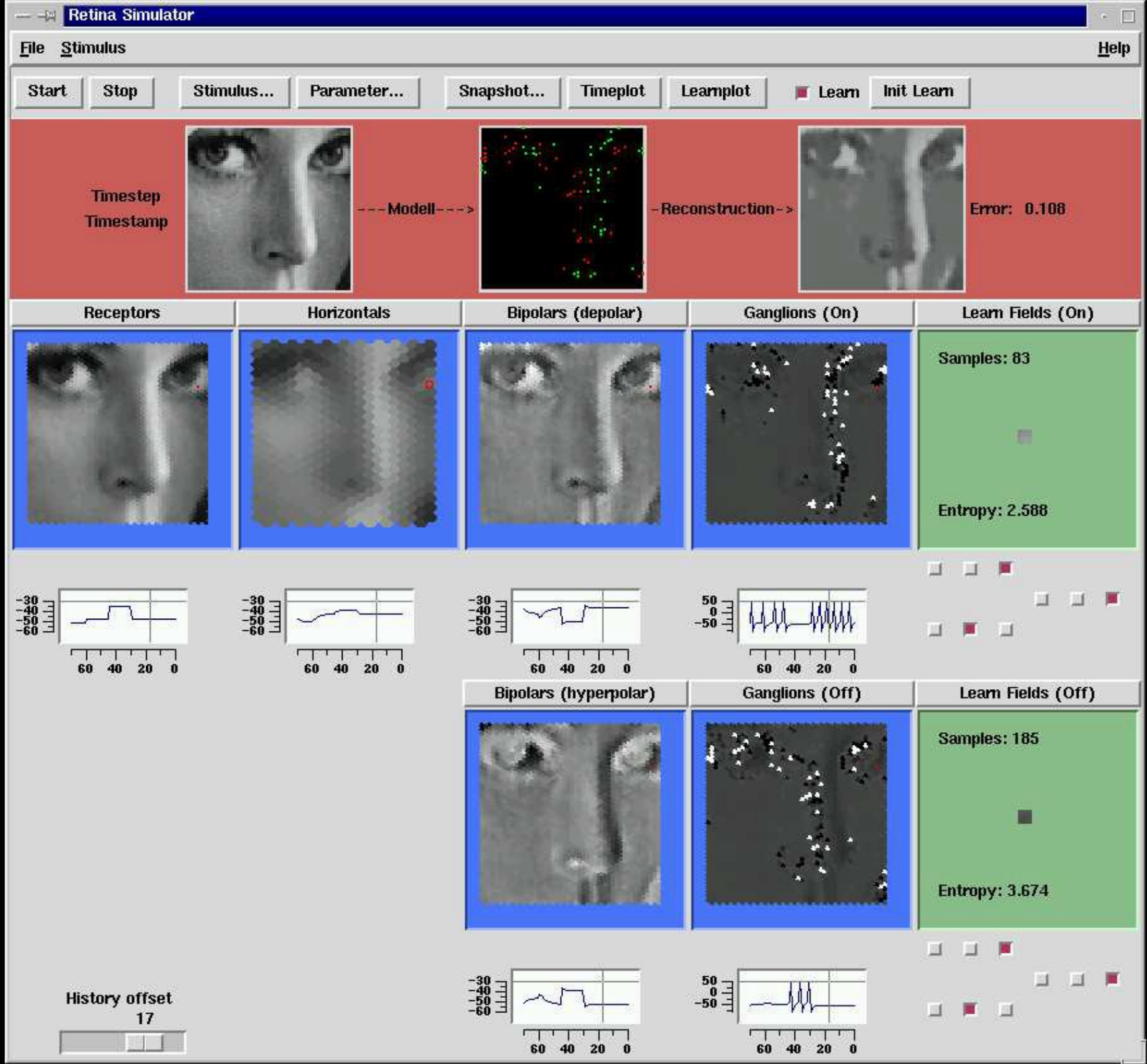
Retinal Stimulus Reconstruction

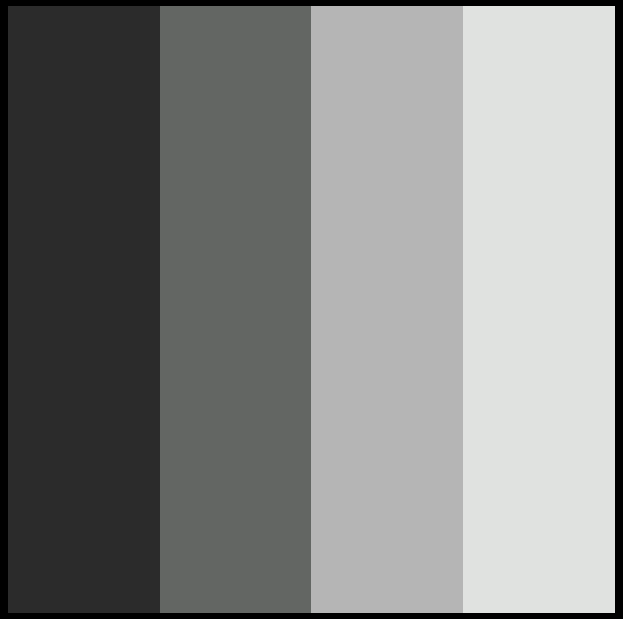


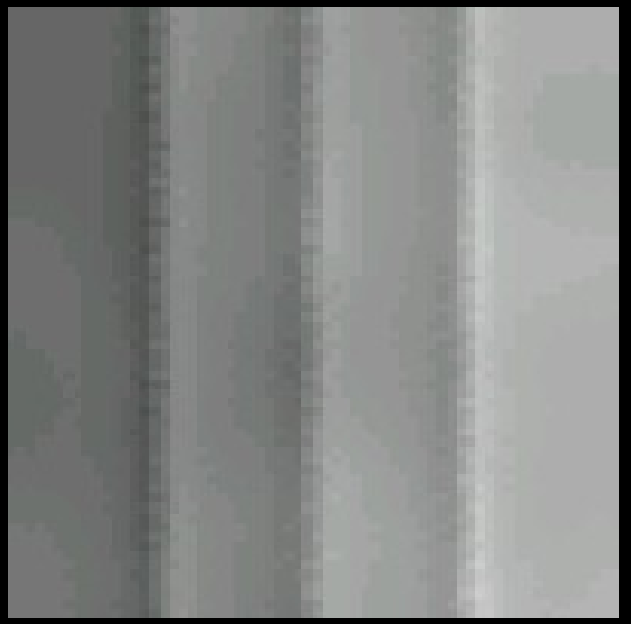


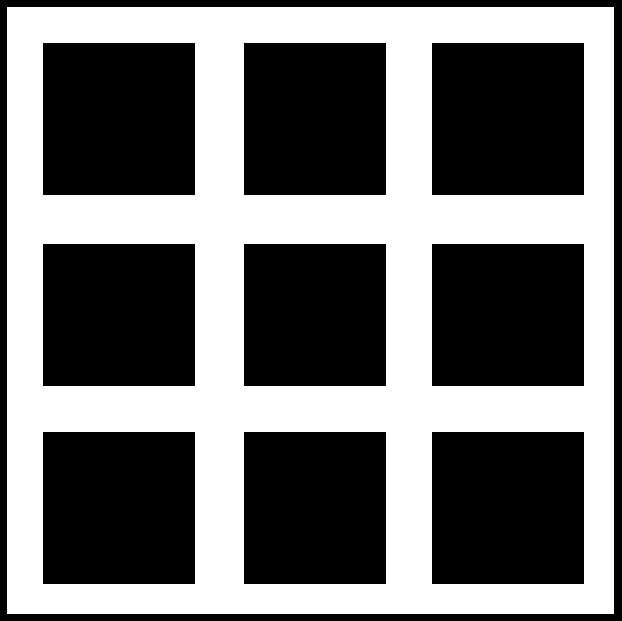
Retinal Stimulus Reconstruction

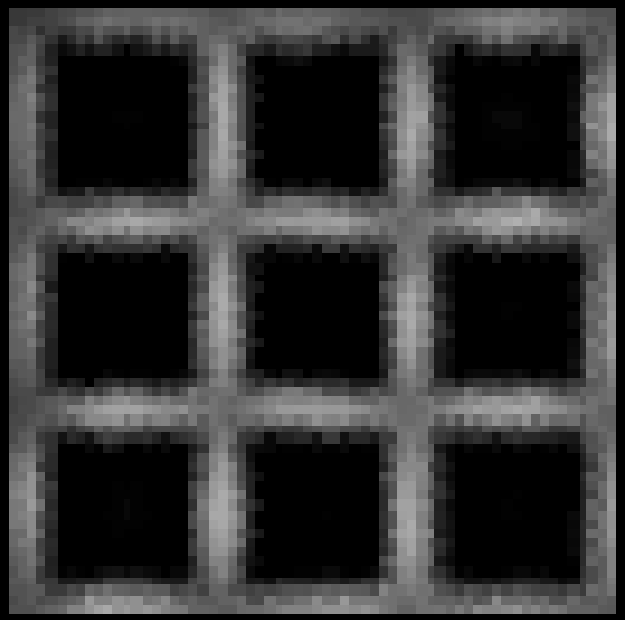


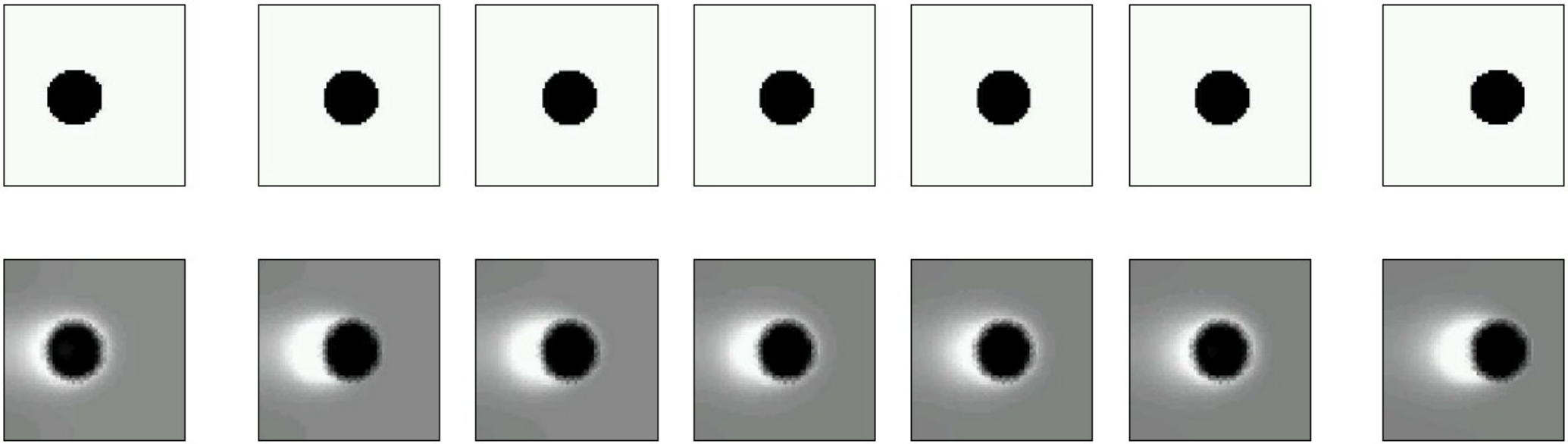

Retina Simulator
File
Stimulus
Help
Start
Stop
Stimulus...
Parameter...
Snapshot...
Timeplot
Learnplot
Learn
Init Learn
Timestep
Timestamp
---Modell--->
-Reconstruction->
Error: 0.108
Receptors
Horizontals
Bipolars (depolar)
Ganglions (On)
Learn Fields (On)
Samples: 83
Entropy: 2.588
-30
-40
-50
-60
60 40 20 0
50
0
-50
Bipolars (hyperpolar)
Ganglions (Off)
Learn Fields (Off)
Samples: 185
Entropy: 3.674
History offset
17











III

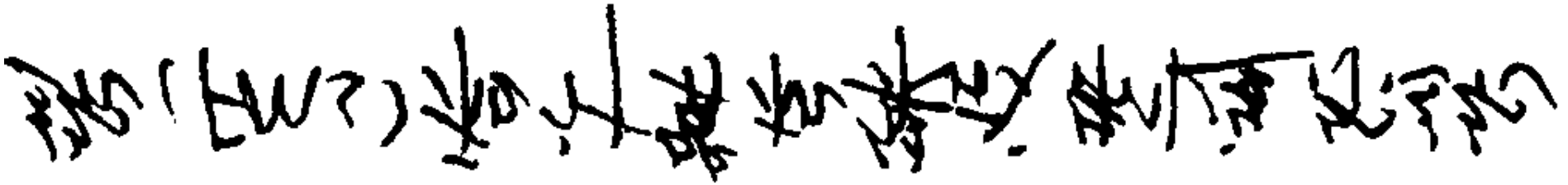
Handwriting Recognition

Hidden Markov Models

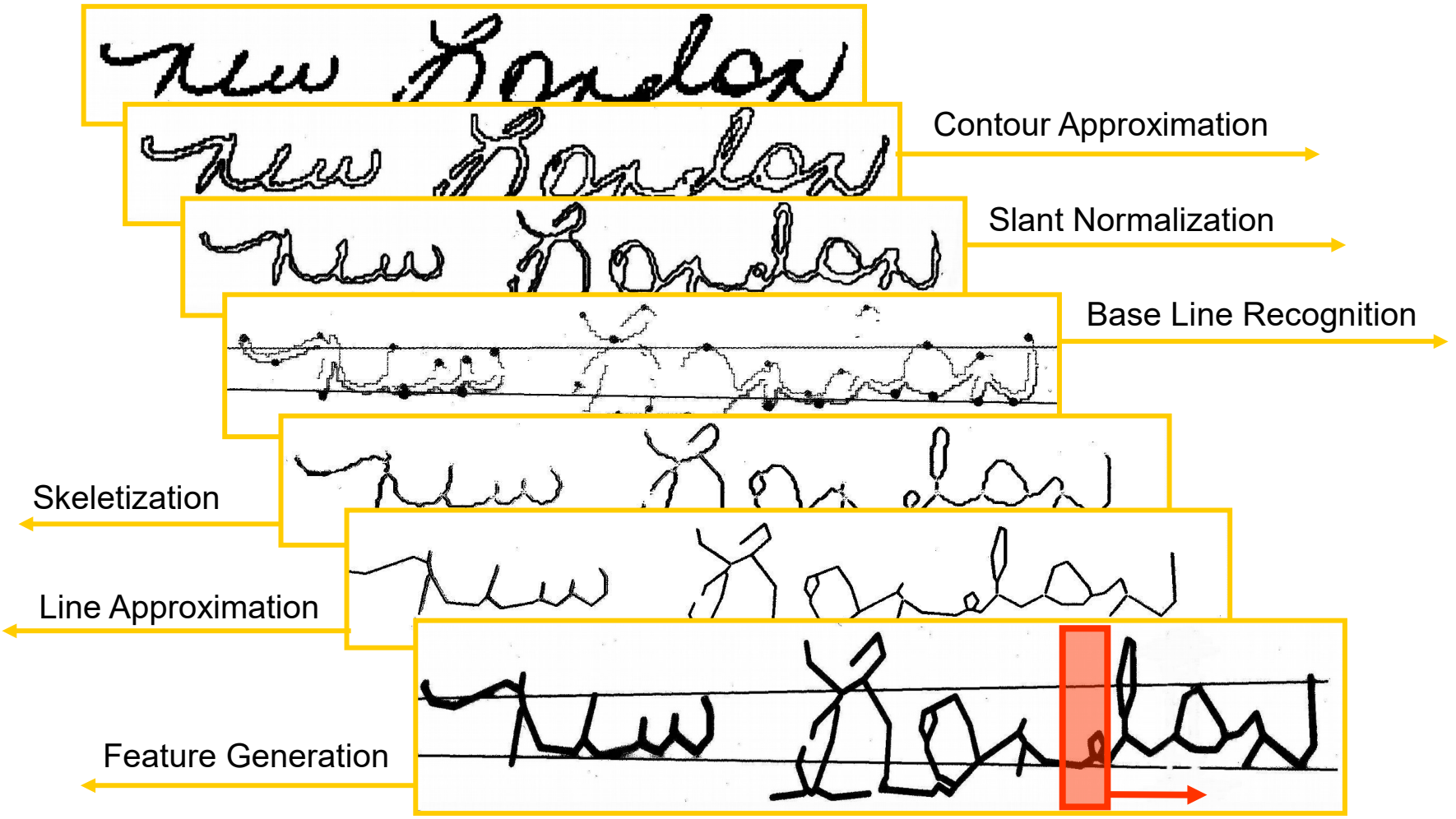

LIGHT IS A
WAVE!

Cursive Script

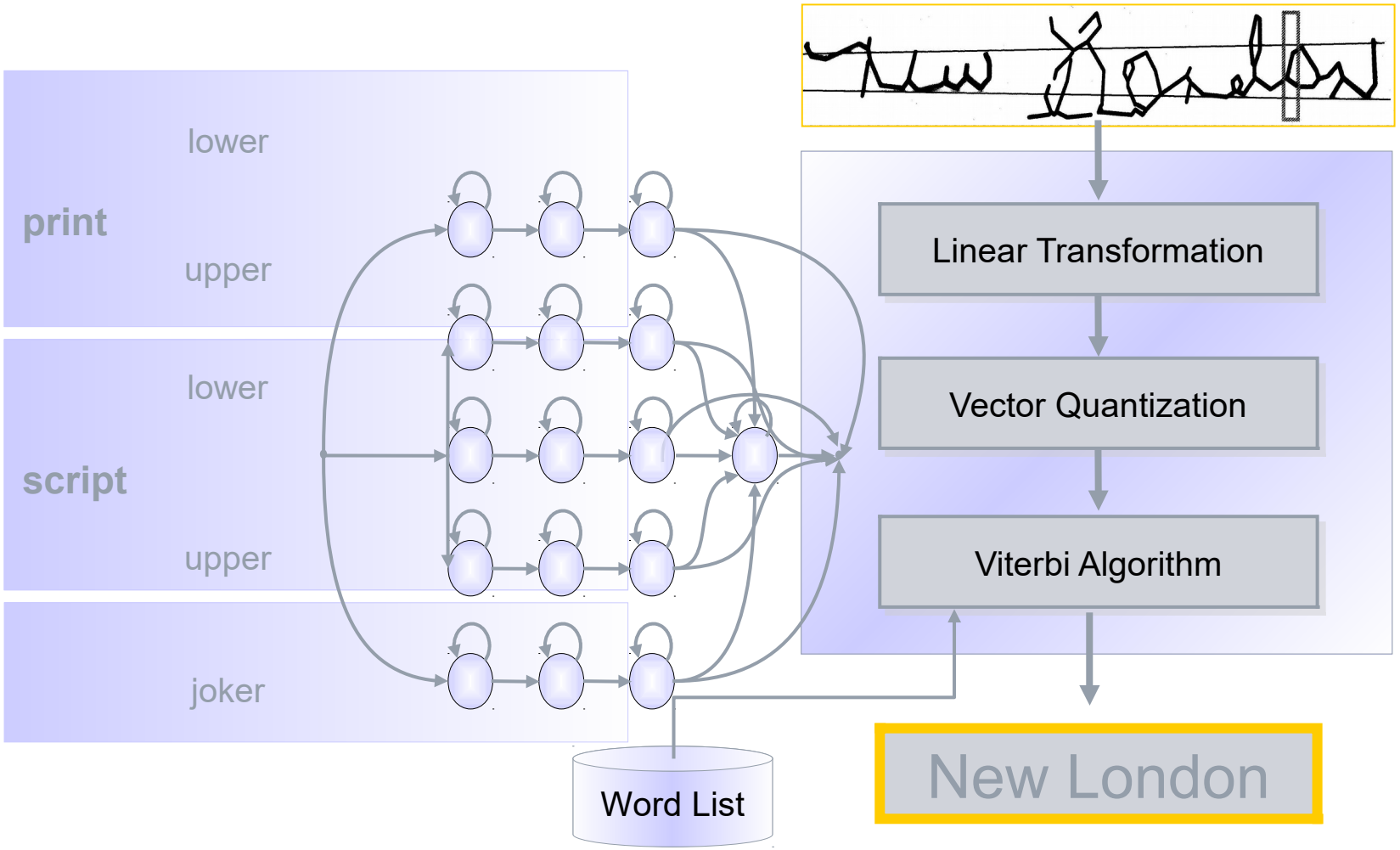
Once having



Preprocessing



Character Models

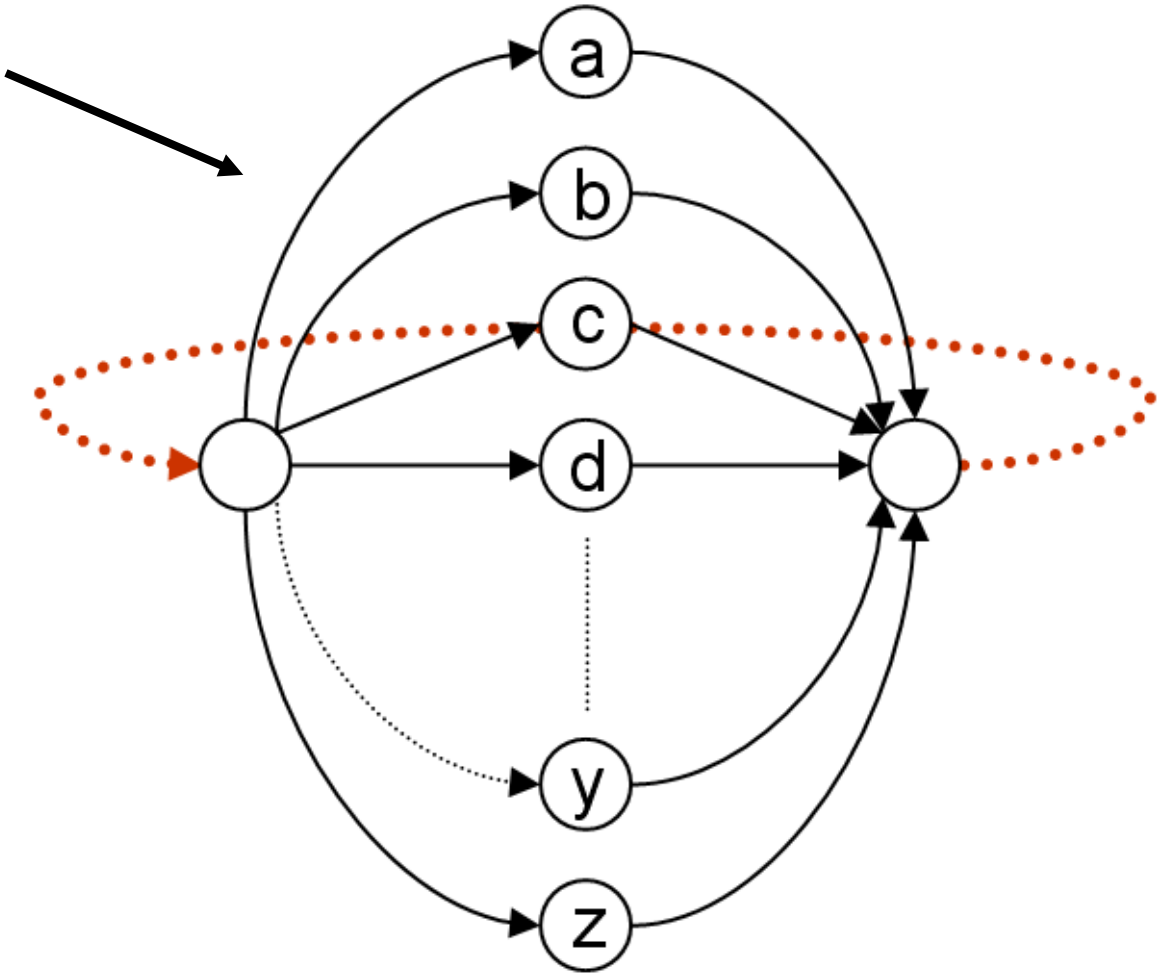


Lexicon-free Recognition

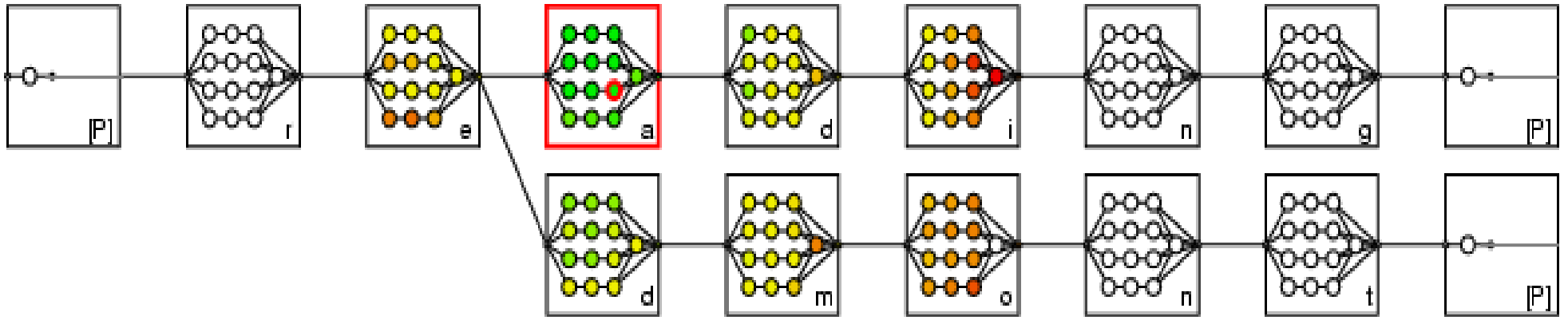
Recurrent HMM

Postprocessing with language models:

- N-grams
- Lexika
- Rules



Viterbi-Decoding



Word models

Prefix-Trie

- Merge word beginnings
- Maximize probability at word endings

Beam-Search

- Calculate only activated nodes
- Deactivate nodes with low likelihood

ICDAR 2007

9th International Conference on Document Analysis and Recognition

September 23-26 2007, Curitiba, Parana, Brazil

Winning Team of the Arabic Handwriting Competition

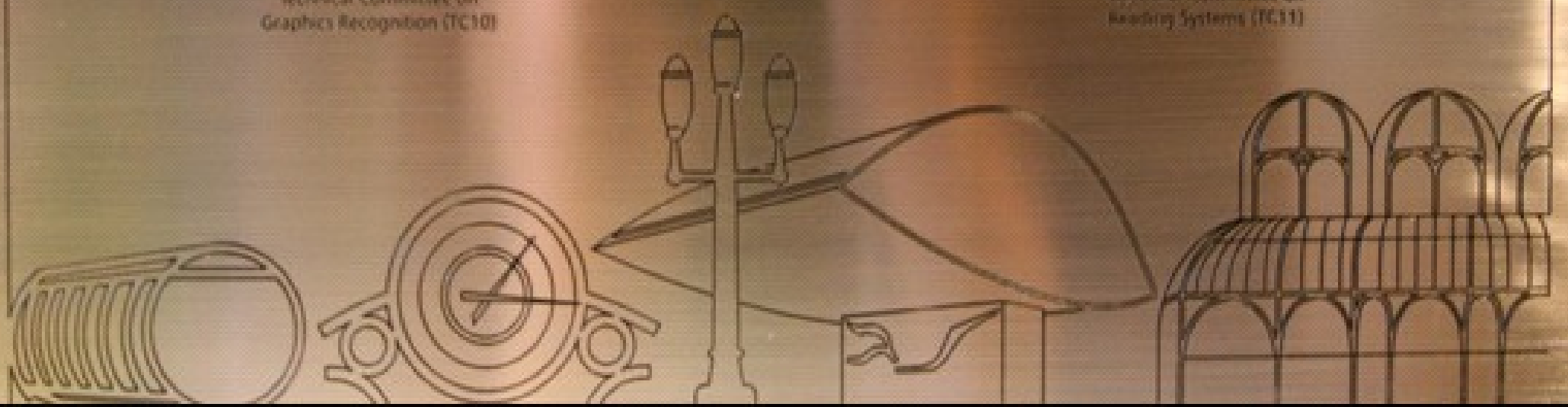
Theophile Alary, Jorg Rottland and Marc-Peter Schambach
Siemens AG Industrial Solutions and Services
Konstanz, Germany

Curitiba, September 2007

IAPR

Technical Committee on Graphics Recognition (TC10)

Technical Committee on Reading Systems (TC11)

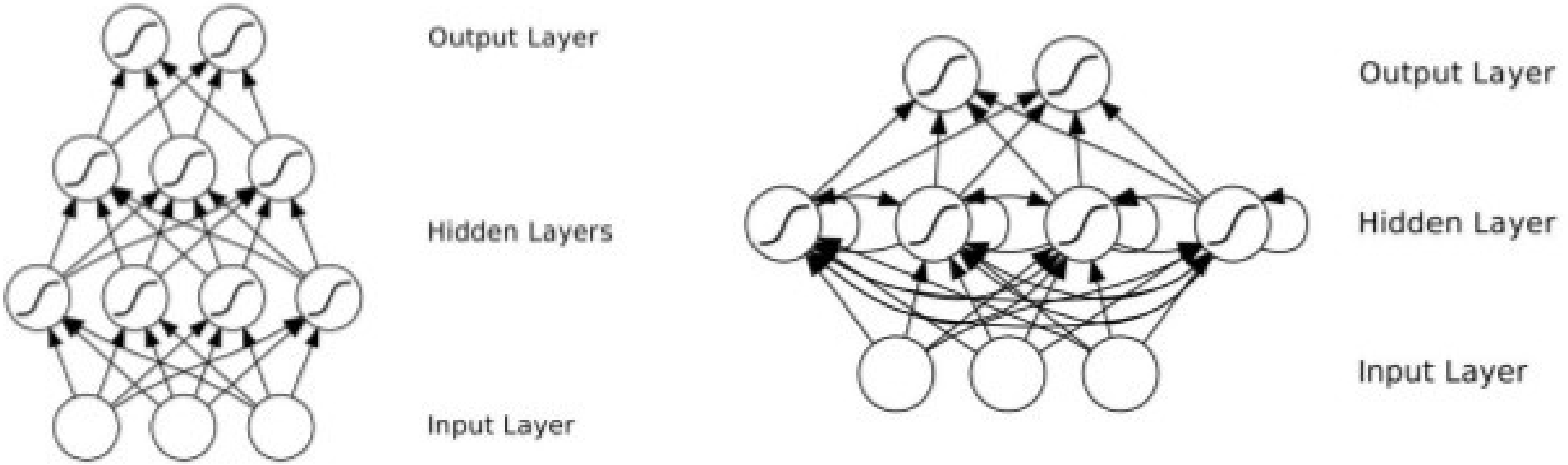


III

Handwriting Recognition

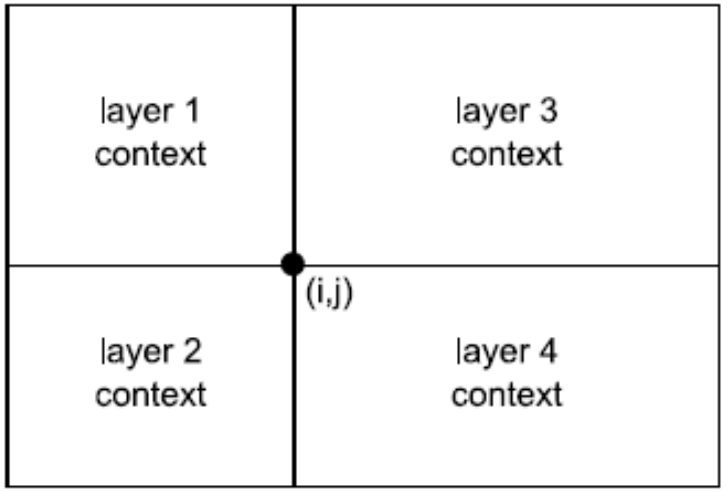
Recurrent Neural Networks

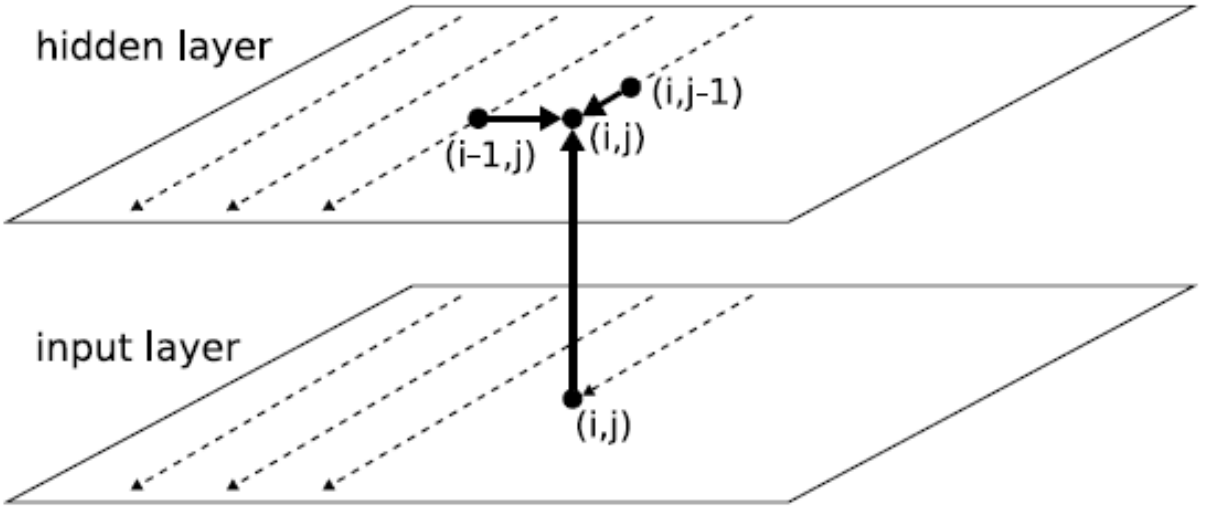
Feedforward NN vs. Recurrent NN

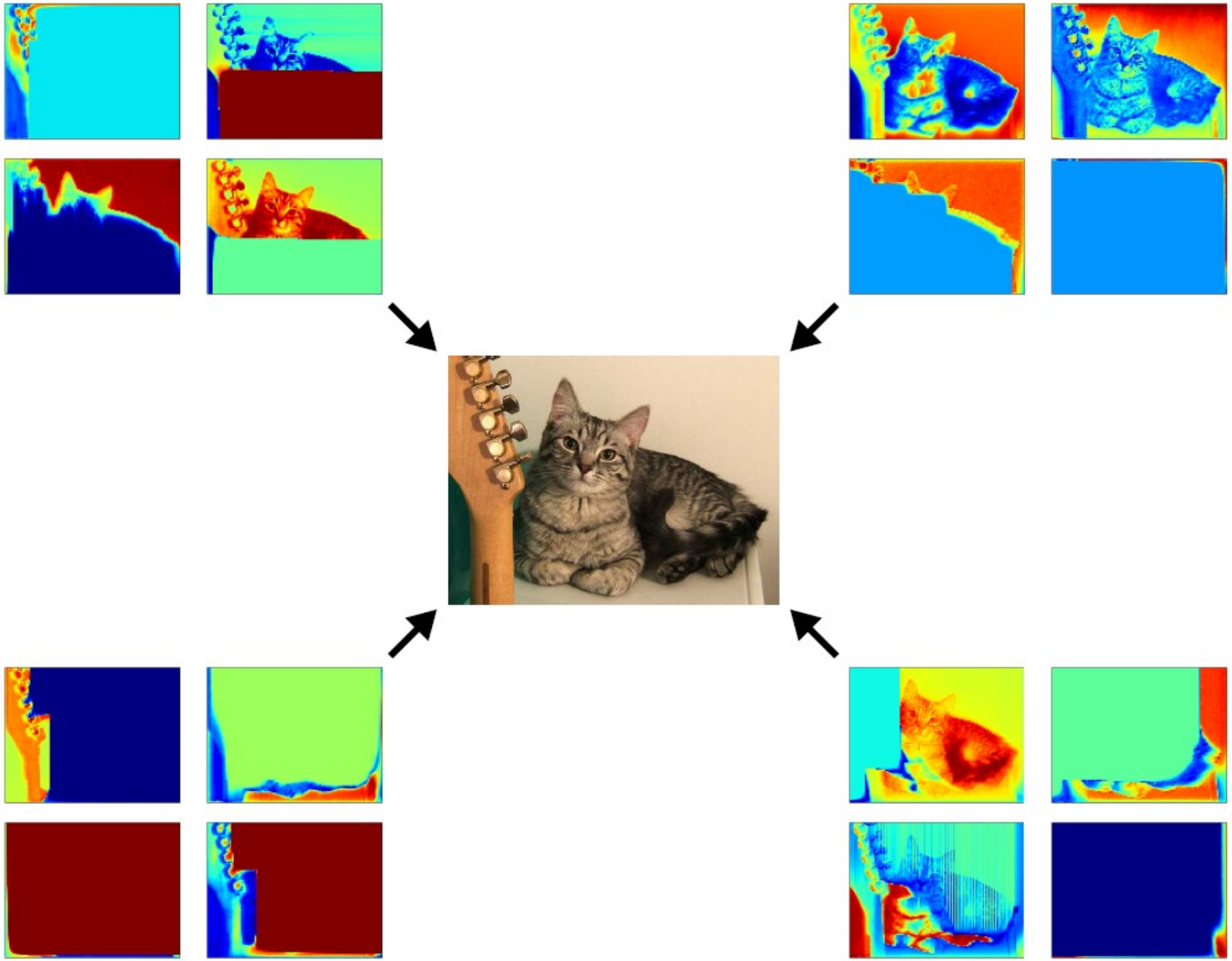


Recurrent neural networks (RNNs) allow cyclical connections.

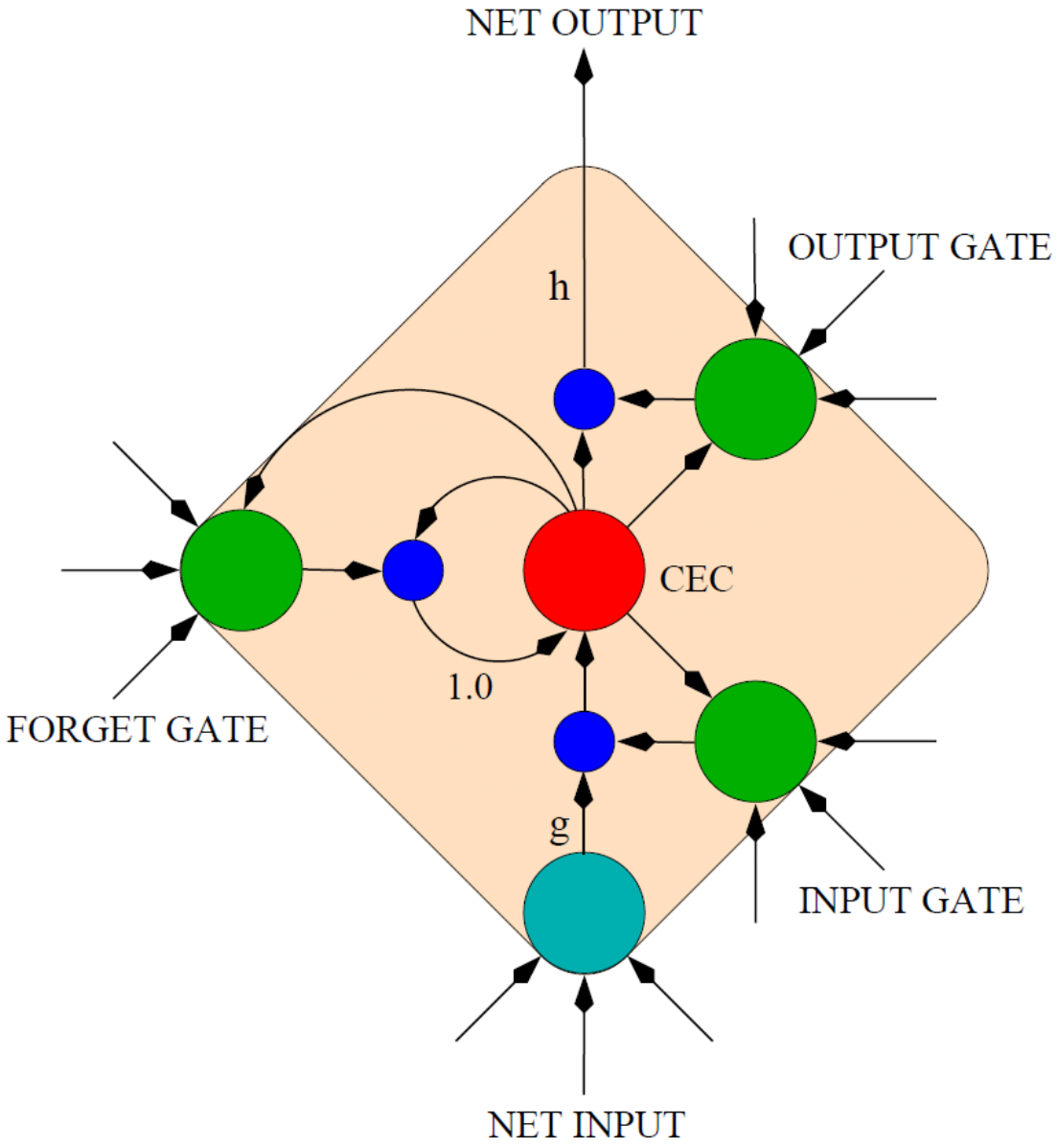
Recurrent Neural Networks



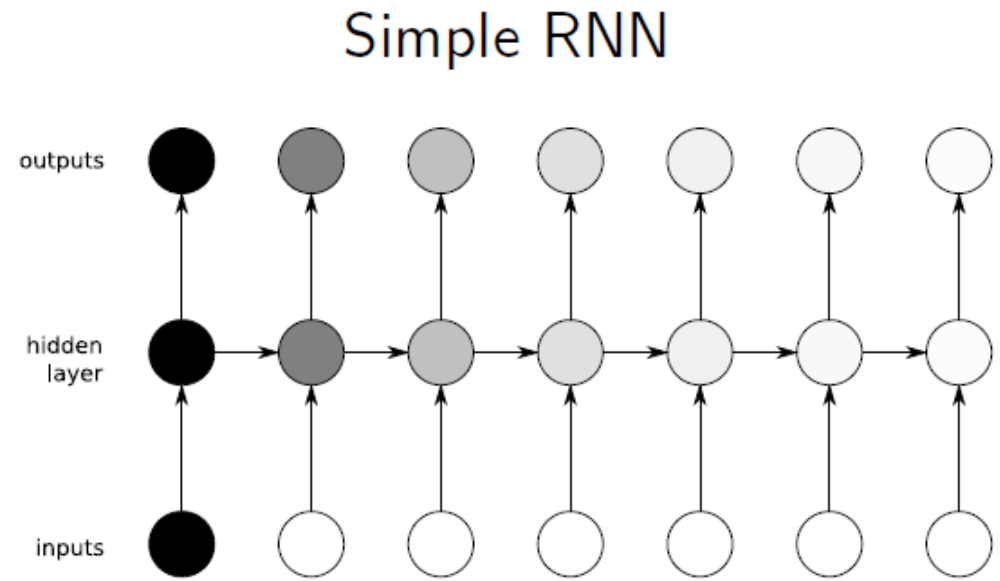


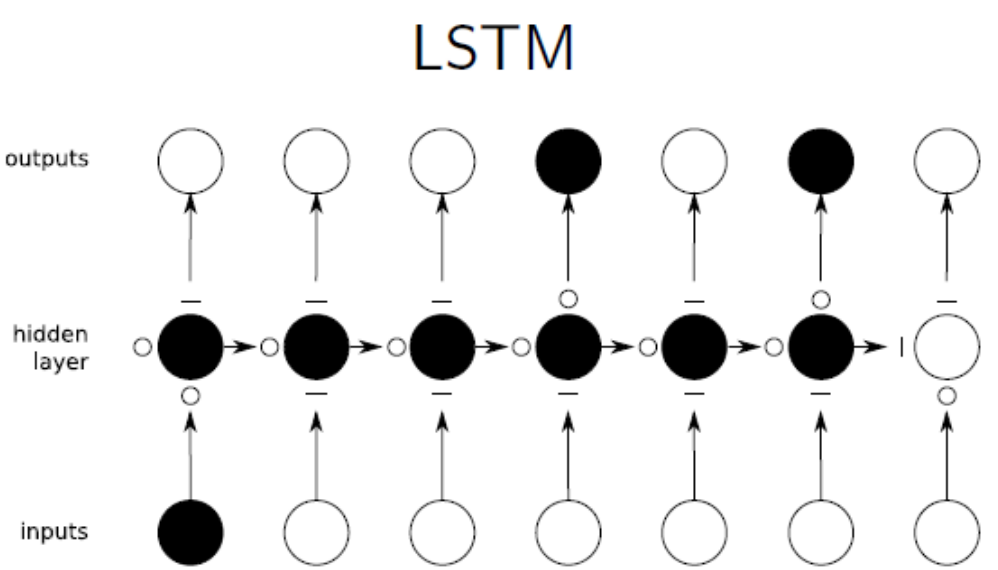


Long short term memory

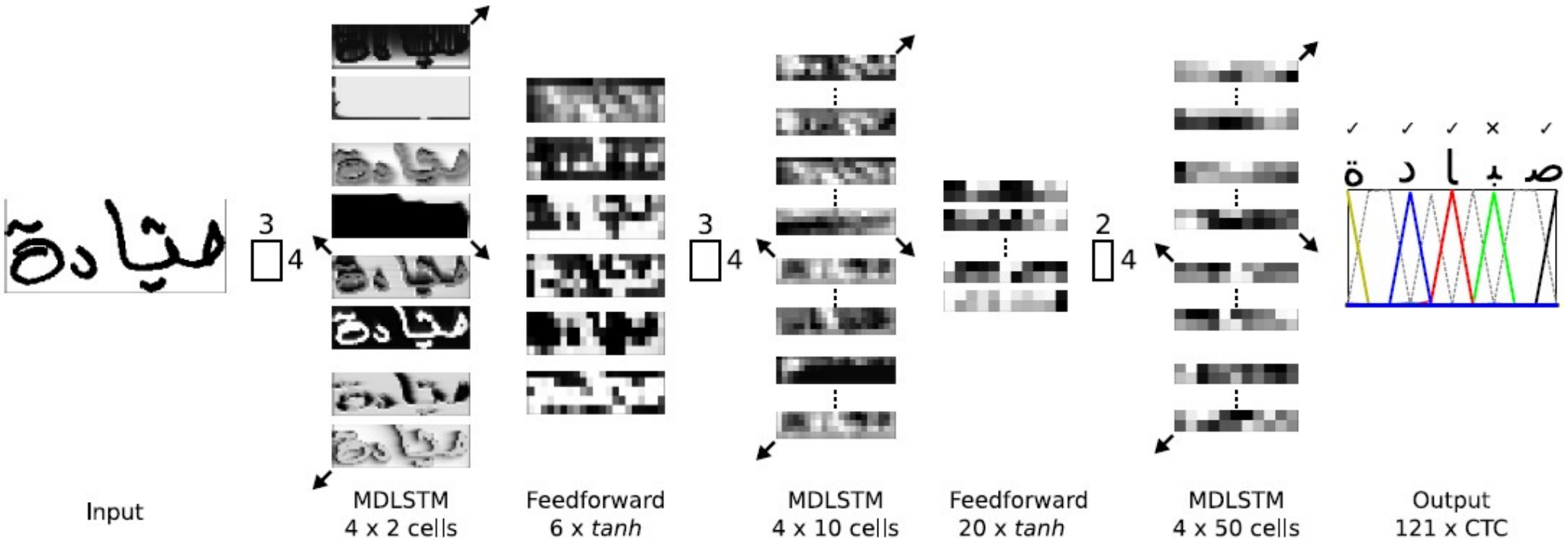


Vanishing Gradient





Multi-dimensional Recurrent Neural Networks

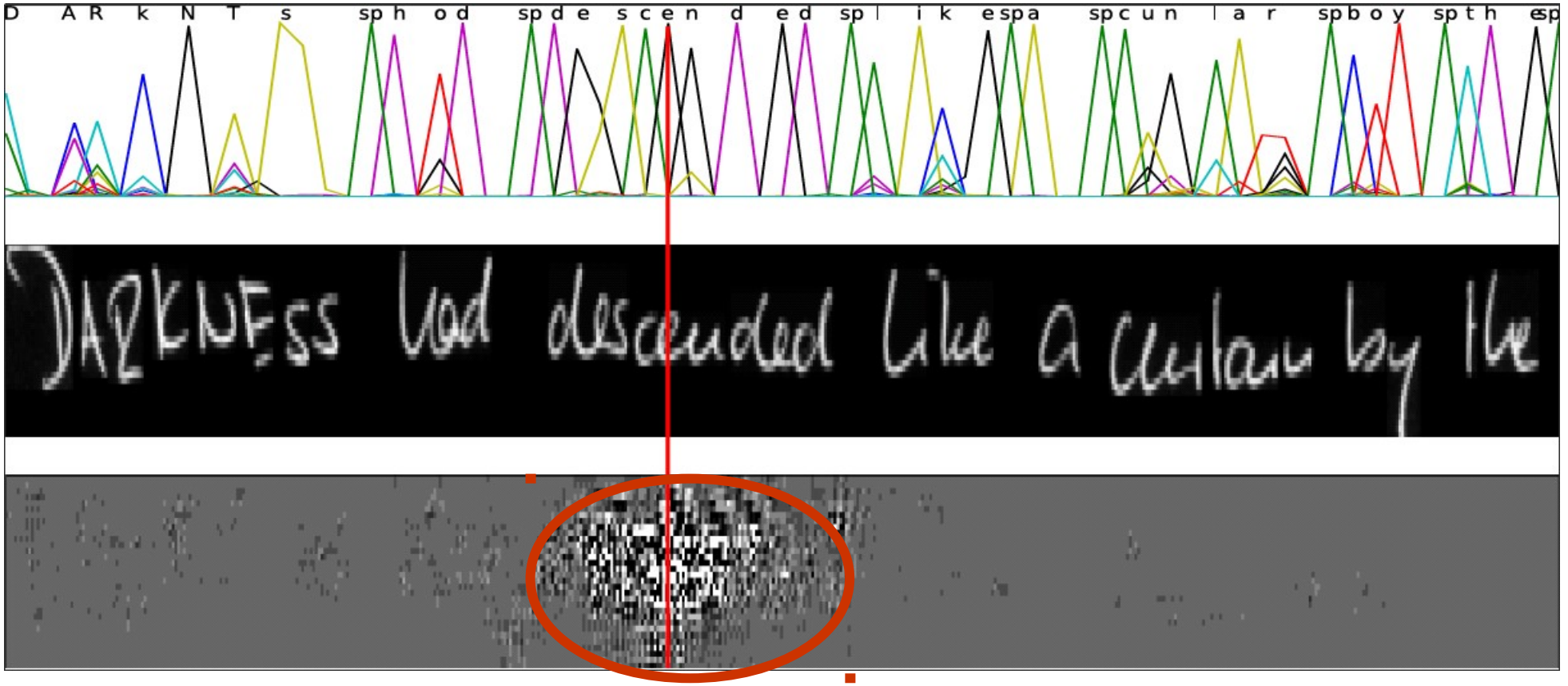


RNN Training

- Backpropagation
- CTC-Layer: Forward-Backward-Alignment determines error function
- Similar to Expectation-Maximization

- Stabilizing by fixed alignment during first training epochs

Recurrent Neural Networks



Jacobi-Matrix: Shows the input range relevant for classification

Handwriting recognition with RNN

+ WRR cursive handwriting 95% (lexicon 10k)
+ CRR machine print 99,6%
+ Robust
+ Trains language models

- Slower than HMM (training and recognition)

International Association for Pattern Recognition

The organizing committee of ICDAR 2013 certifies that

Marc-Peter Schambach, Joerg Rottland, and Sheikh Faisal Rashid

won First Place in the

ICDAR 2013 Competition on Multi-font and Multi-size Digitally Represented Arabic Text

August 28th, 2013

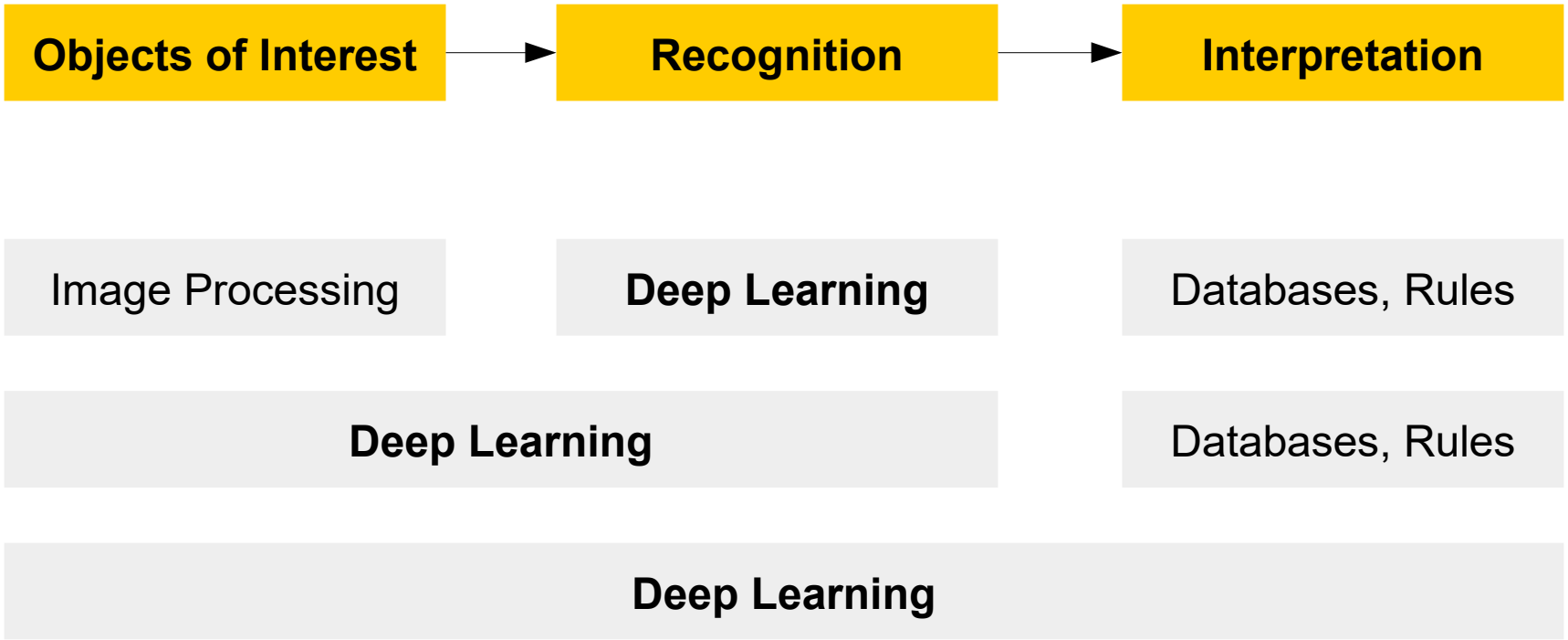
ICDAR 2013 General Chair

III

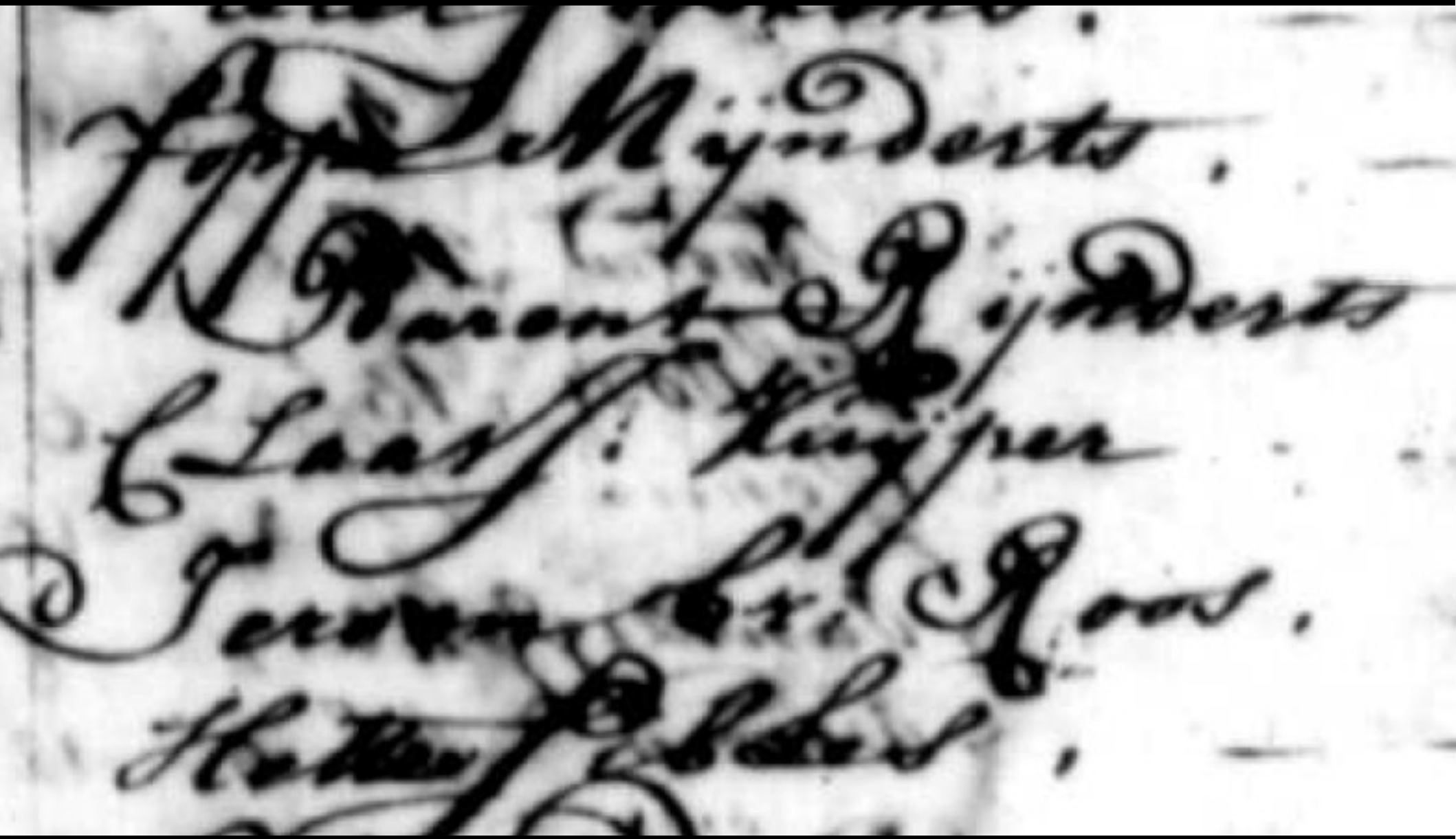
Reading & Interpretation
Vision & Understanding

Future Directions

Vision & Understanding: Operation Pipeline



Simon Geerts, –
Douwe Tjamkes. –
Marten Buurs. –
Arnolt de Haan, –
Dirk Telles. –



IV

Deep Learning

Recent Applications

Photo Geolocation

with

Convolutional Neural Networks

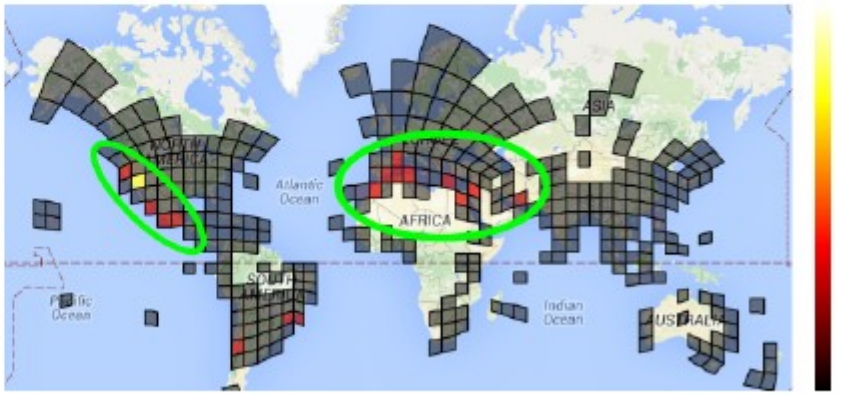
Photo CC-BY-NC by stevekc

(a)

Photo CC-BY-NC by edwin.11

(b)

Photo CC-BY-NC by jonathanfh



(c)

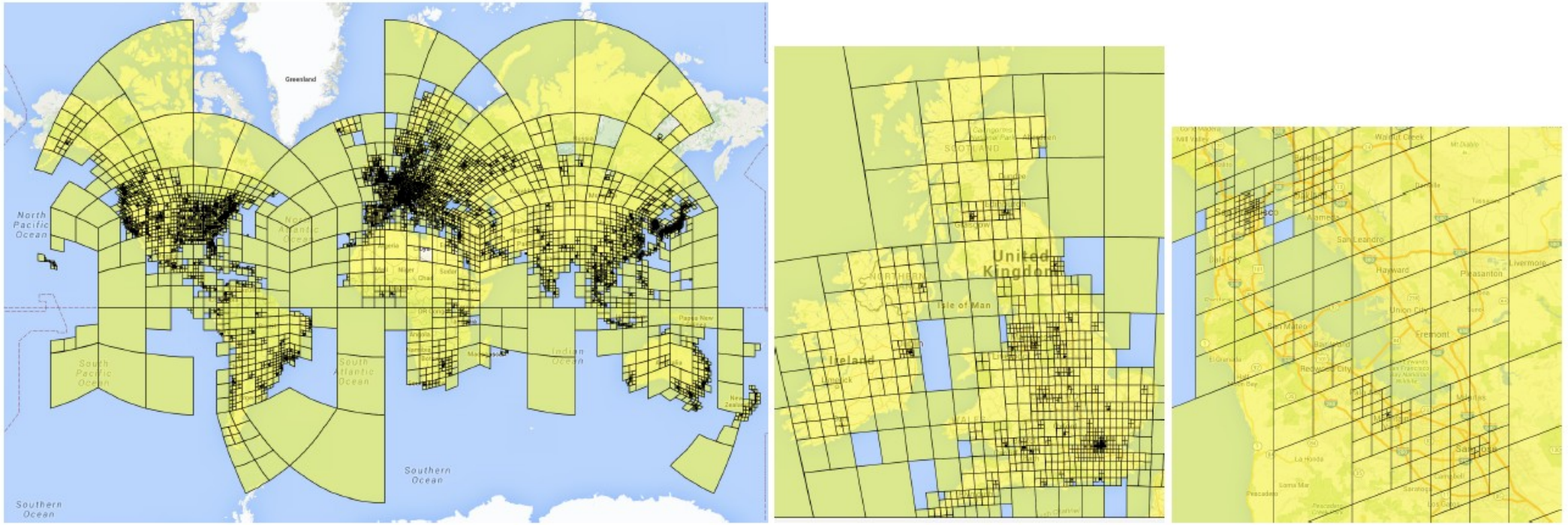
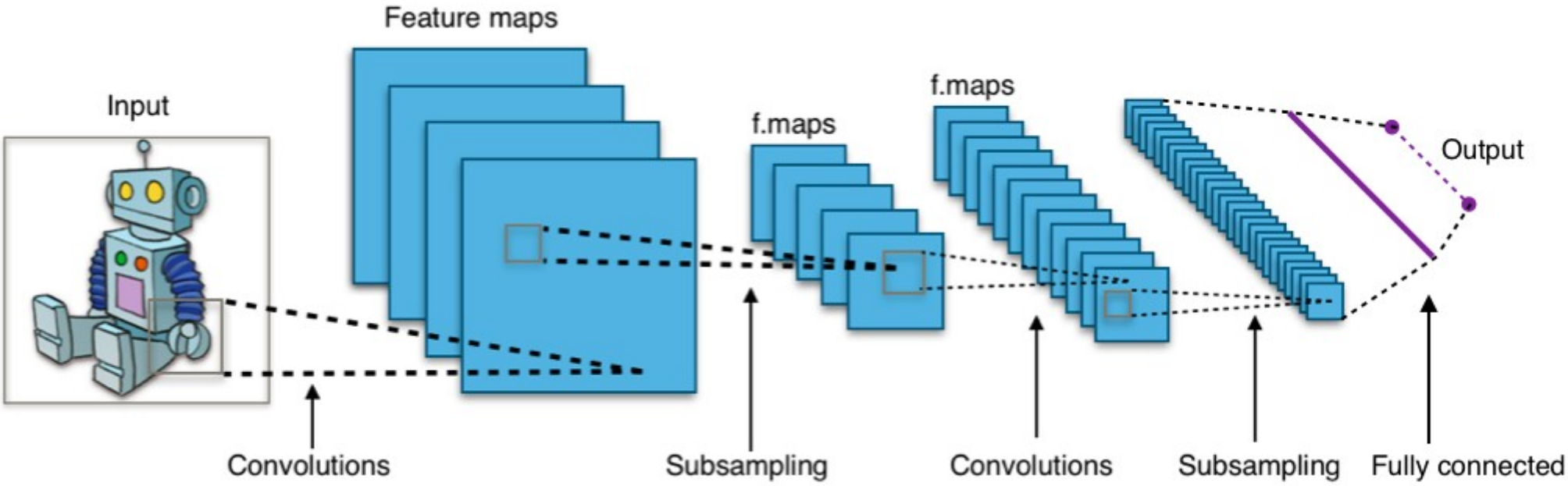


Figure 2. Left: Adaptive partitioning of the world into 26,263 S2 cells. Right: Detail views of Great Britain and Ireland and the San Francisco bay area.

Convolutional Neural Networks



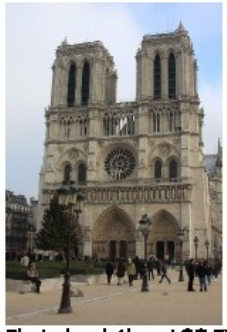


Photo by dotben / CC BY NC

Paris, France

Photo by simononly / CC BY NC

London, England

Photo by judepics / CC BY NC

New York, USA

Photo by hannah takes pictures / CC BY NC

Sydney, Australia

Photo by kenner116 / CC BY NC

Beijing, China

Photo by iogi / CC BY NC

Tokyo, Japan

Photo by palebluejen / CC BY NC

Beijing, China

Photo by sneakysdream / CC BY NC

London, England

Photo by nehavish / CC BY NC

Paris, France

Photo by Bluedevil266 / CC BY NC

Los Angeles, USA

Photo by Duarte Antunes / CC BY NC

Iceland

Photo by habi / CC BY NC

Iceland

Photo by Loimere / CC BY NC

Northern Scotland

Photo by dalbera / CC BY NC

Northern India

Photo by Rex Gray / CC BY NC

Pred: Hanava, Cuba
GT: Long Beach, USA

Photo by ChrisL_AK / CC BY NC

Pred: Iceland
GT: Alaska, USA

Photo by J. Stephen Conn / CC BY NC

Pred: Scotland
GT: Alaska, USA

Photo by Tomcio77 / CC BY NC

Pred: Seychelles
GT: British Virgin Islands

Namibia / Botswana

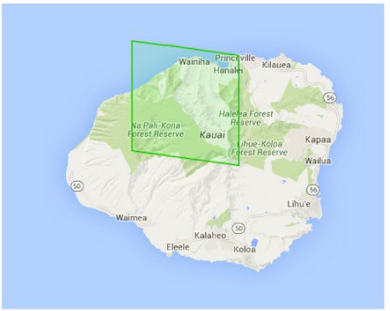
Photo by jamie.lovelock / CC BY NC

Photo by MongoosePhotography / CC BY NC

Photo by Mister-E / CC BY NC

Photo by dalangalma / CC BY NC

Photo by siamjack / CC BY NC



Kauai, Hawaii

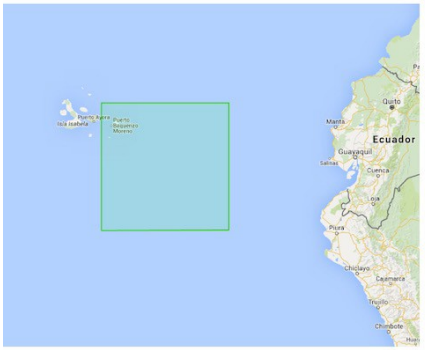
Photo by ryan + sarah / CC BY NC

Photo by stuartlchambers / CC BY NC

Photo by samgrover / CC BY NC

Photo by steuben / CC BY NC

Photo by steve-stevens / CC BY NC



Galapagos Islands

Photo by p.j.k. / CC BY NC

Photo by victor408 / CC BY NC

Photo by Domen Jakus / CC BY NC

Photo by cvanholder / CC BY NC

Photo by rwoan / CC BY NC

Paris

Photo by feliven / CC BY NC

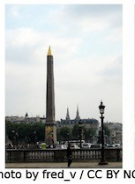


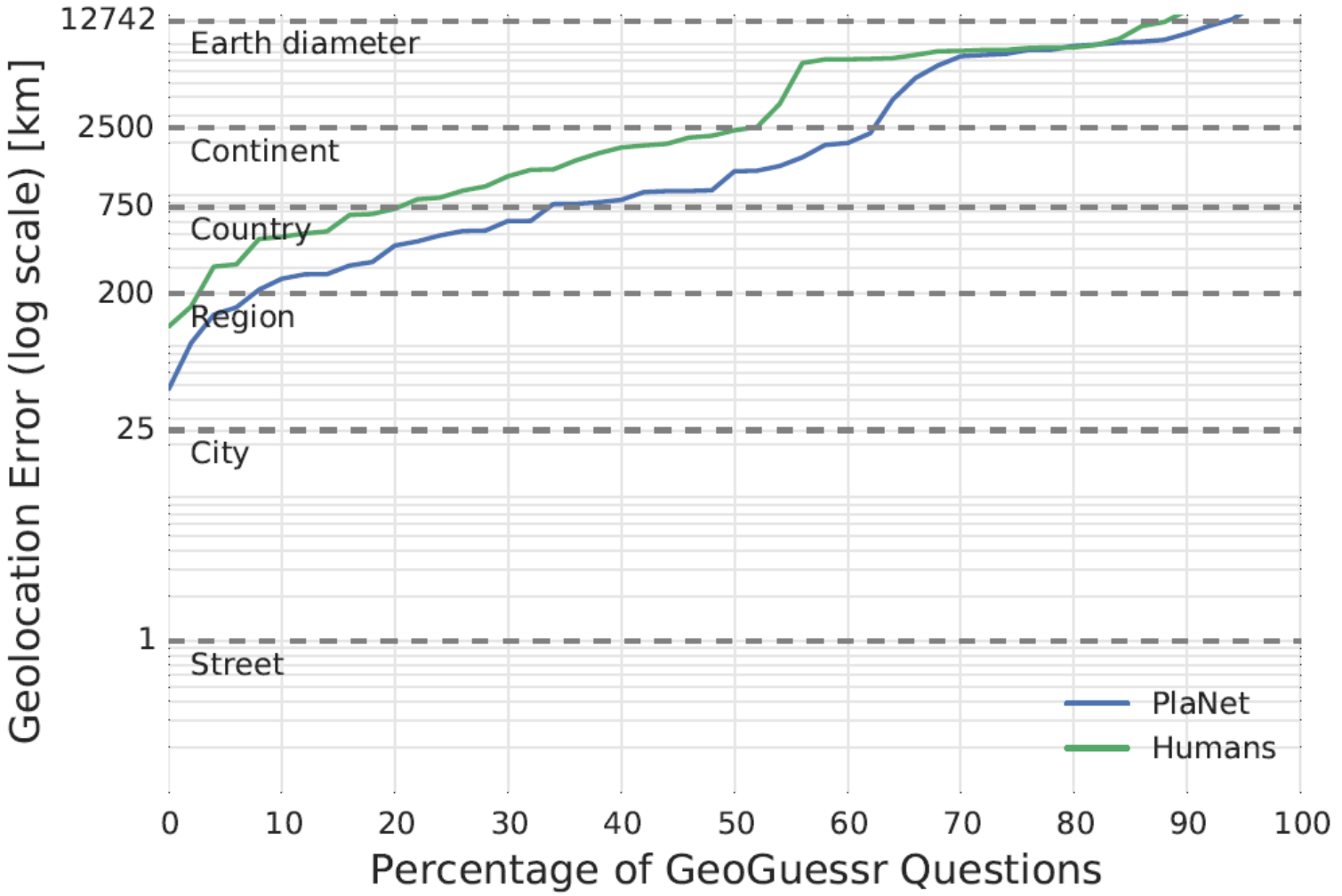
Photo by fred_v / CC BY NC

Photo by Turansa Tours / CC BY NC

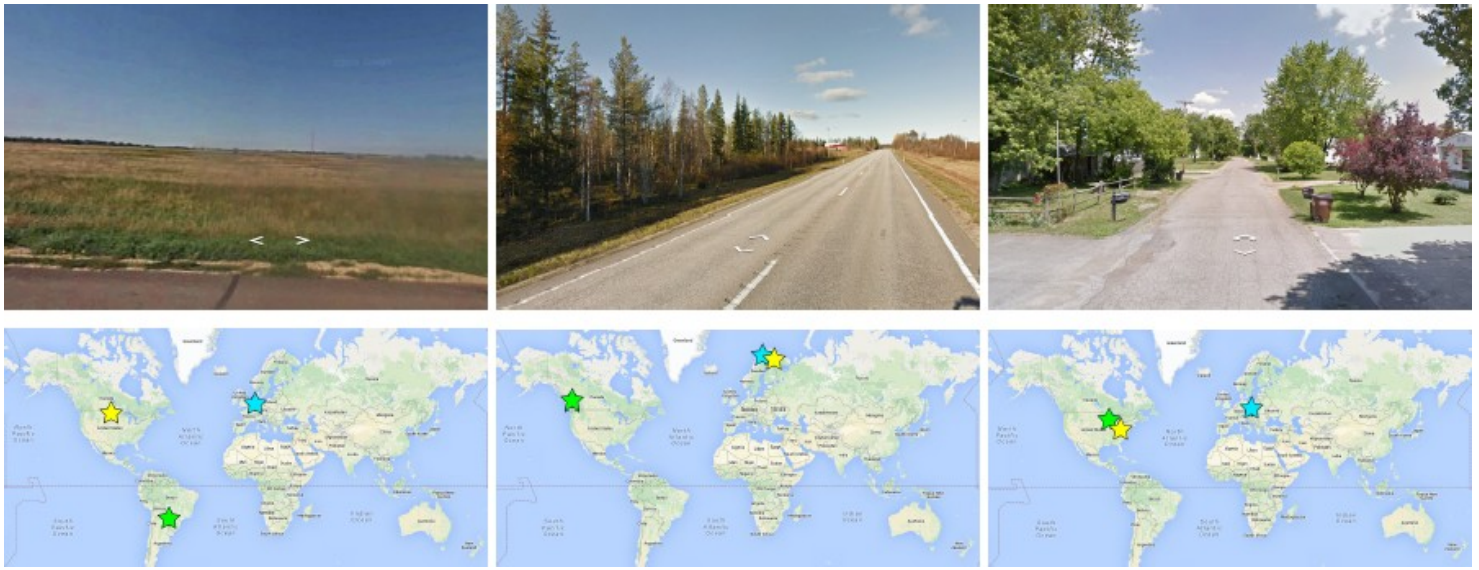
Photo by JA_FS / CC BY NC

Photo by CedEm photographies / CC BY NC

Geolocation Error: PlaNet vs. Humans



More examples



GeoGuessr Panorama

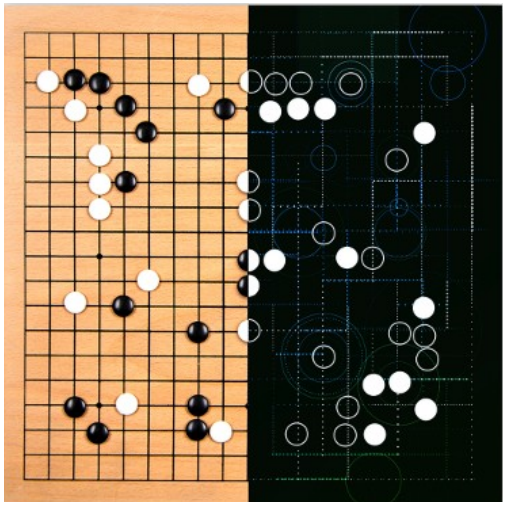
Ground truth

Human guess

PlaNet guess

Mastering the Game of Go with Deep Neural Networks and Tree Search

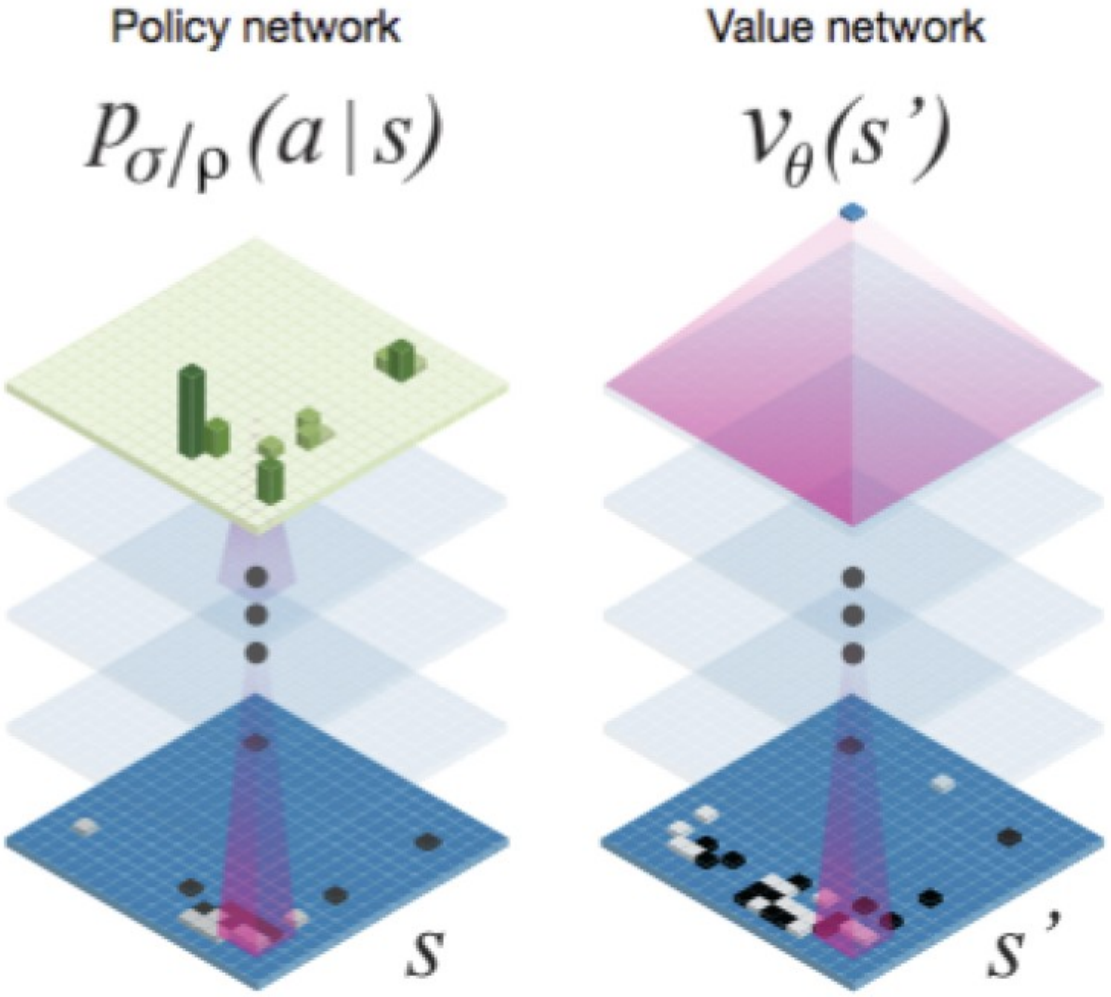
AlphaGo



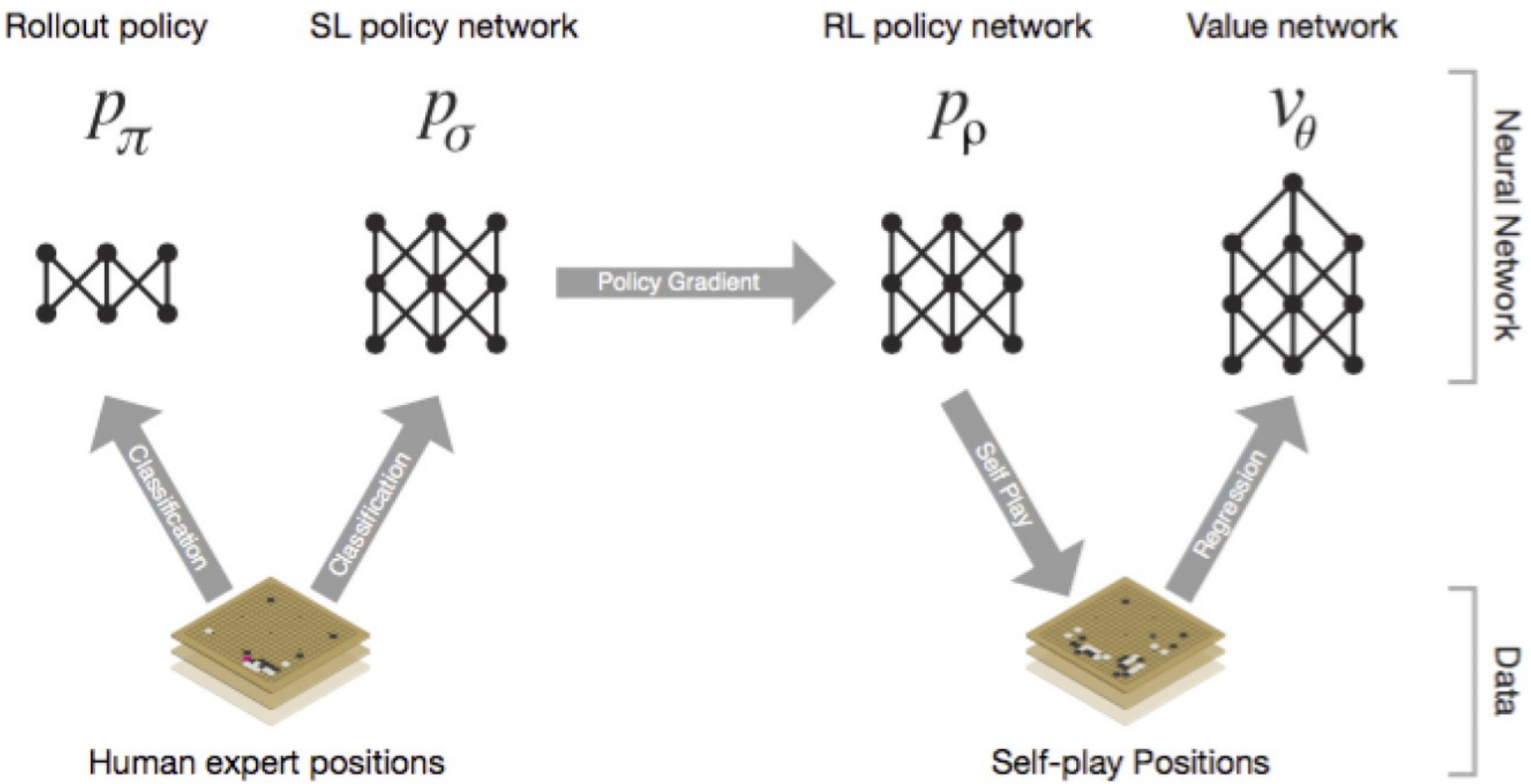
December 2015:
AlphaGo defeats Fan Hui 5–0

March 2016:
AlphaGo defeats Lee Sedol 4–1

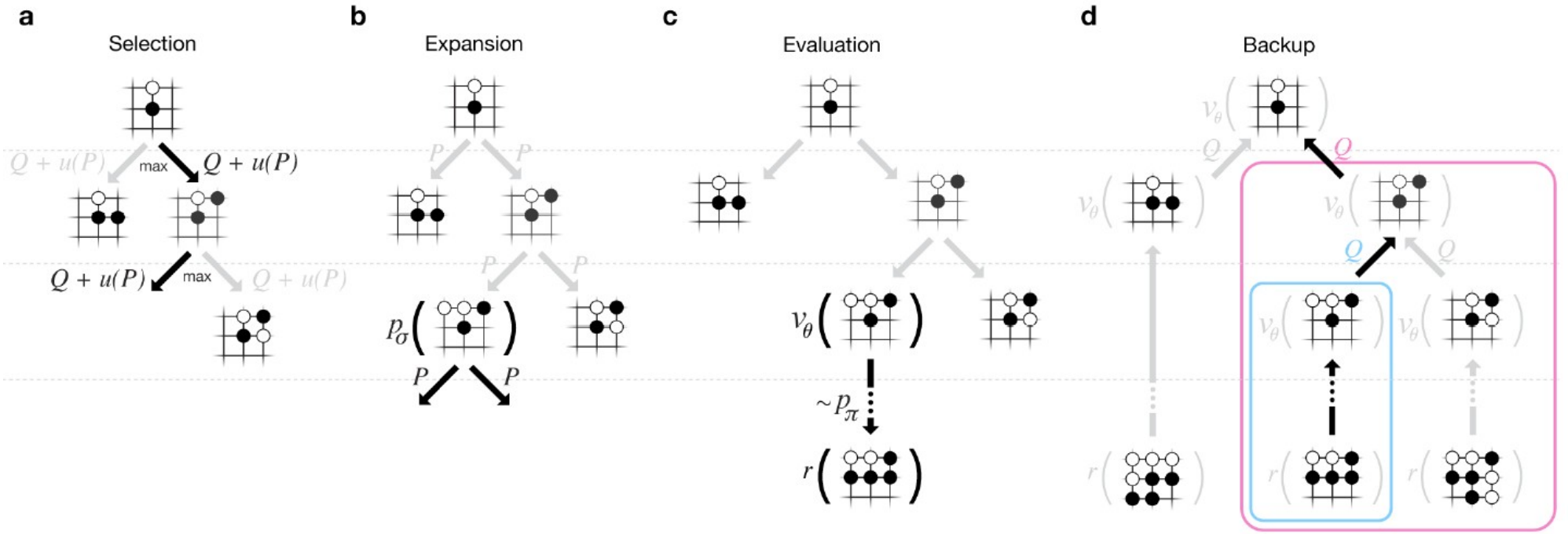
Neural Network Architecture



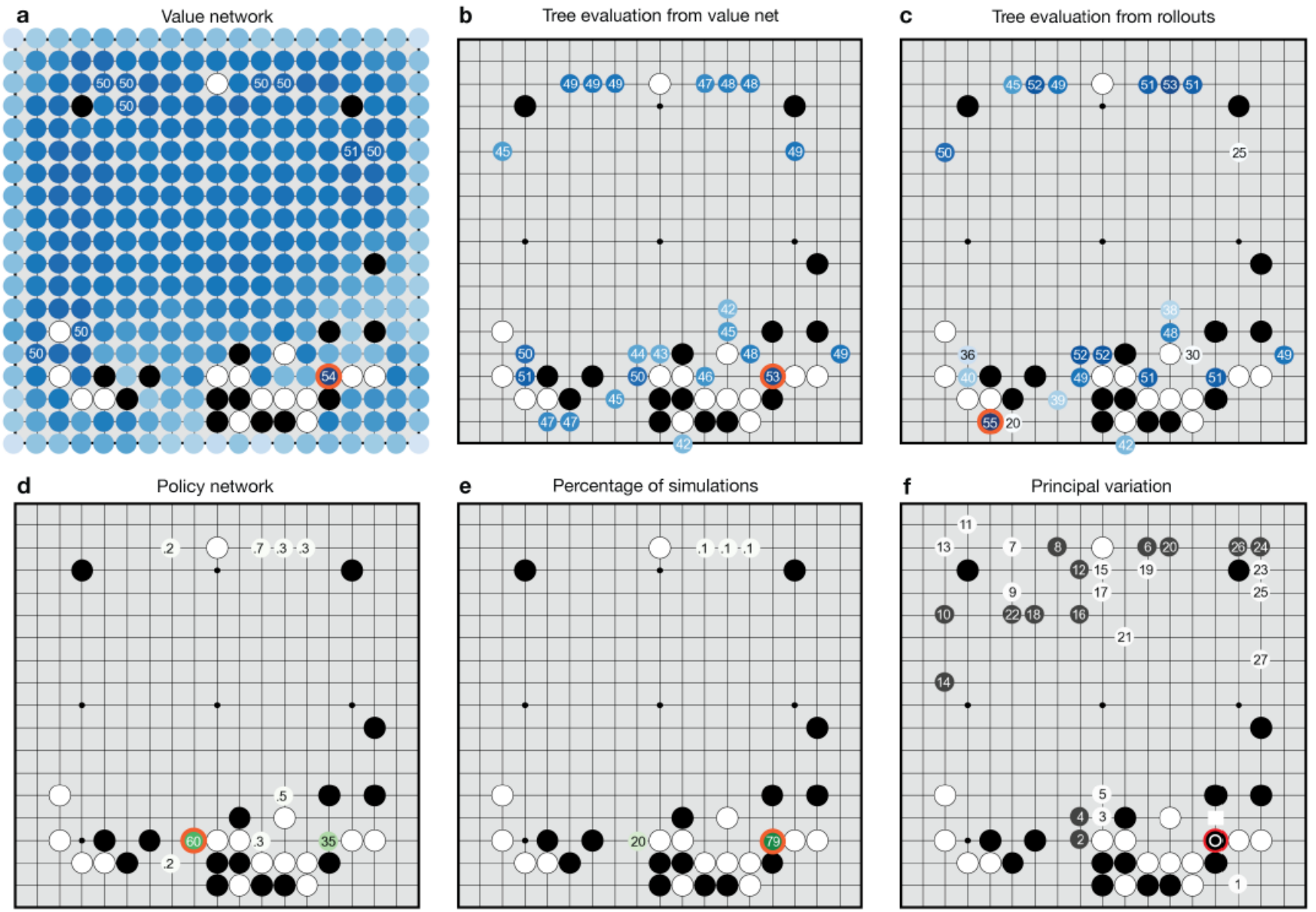
Neural Network Training Pipeline



Monte Carlo Tree Search



Selecting Moves



AlphaGo: Results

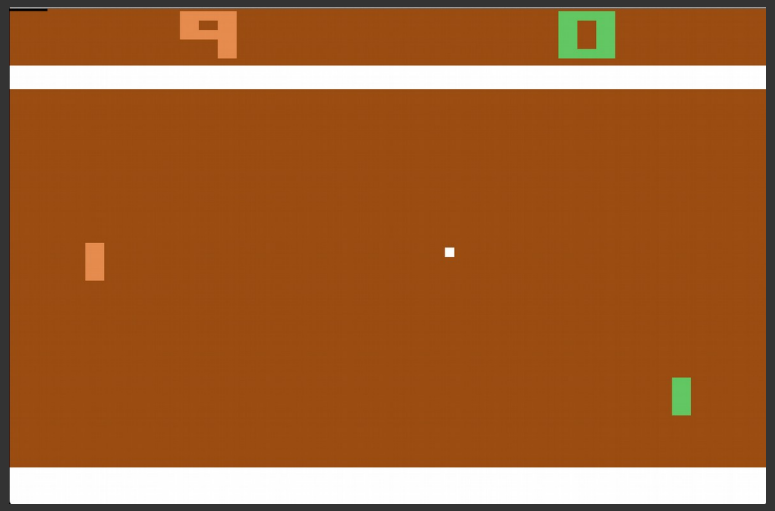
Short name	Computer Player	Version	Time settings	CPUs	GPUs	KGS Rank	Elo
α^d_{rvp}	Distributed *AlphaGo*	See Methods	5 seconds	1202	176	–	3140
α_{rvp}	*AlphaGo*	See Methods	5 seconds	48	8	–	2890
CS	CrazyStone	2015	5 seconds	32	–	6d	1929
ZN	Zen	5	5 seconds	8	–	6d	1888
PC	Pachi	10.99	400,000 sims	16	–	2d	1298
FG	Fuego	svn1989	100,000 sims	16	–	–	1148
GG	GnuGo	3.8	level 10	1	–	5k	431
CS_4	CrazyStone	4 handicap stones	5 seconds	32	–	–	2526
ZN_4	Zen	4 handicap stones	5 seconds	8	–	–	2413
PC_4	Pachi	4 handicap stones	400,000 sims	16	–	–	1756

Extended Data Table 6: **Results of a tournament between different Go programs.** Each program played with a maximum of 5 seconds thinking time per move; the games against Fan Hui were conducted using longer time controls, as described in Methods. CS_4, ZN_4 and PC_4 were given 4 handicap stones; *komi* was 7.5 in all games. Elo ratings were computed by *BayesElo*.

Playing Atari

with

Deep Reinforcement Learning



Pong

Breakout

Space Invaders

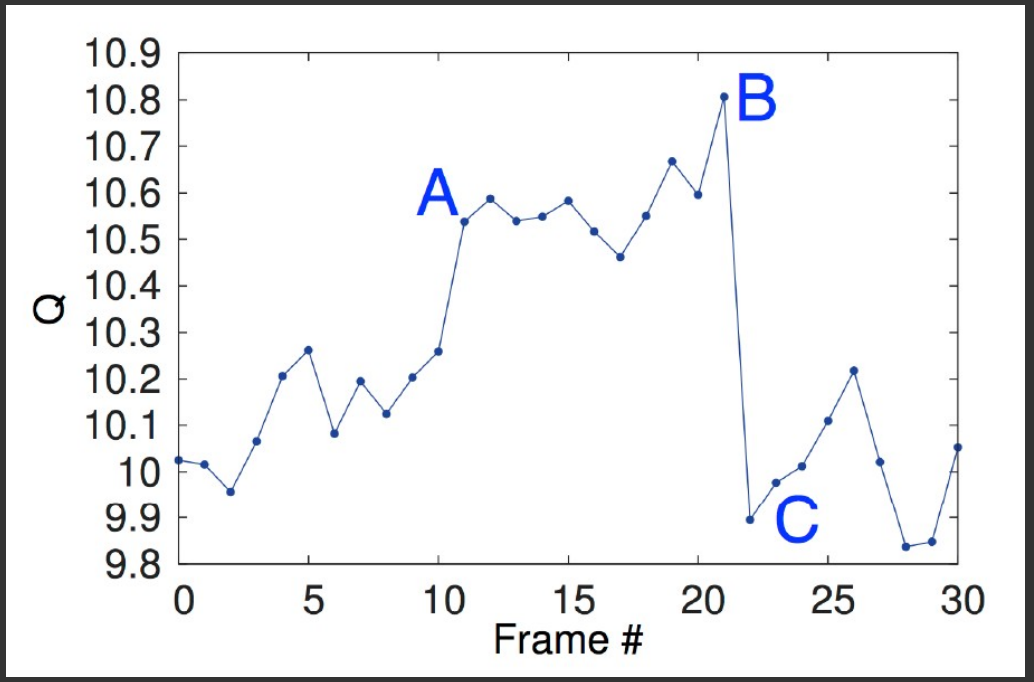
Seaquest

Beam Rider

A

B

C


10.9
10.8
10.7
10.6
10.5
10.4
10.3
10.2
10.1
10
9.9
9.8
Q
0
5
10
15
20
25
30
Frame #
A
B
C

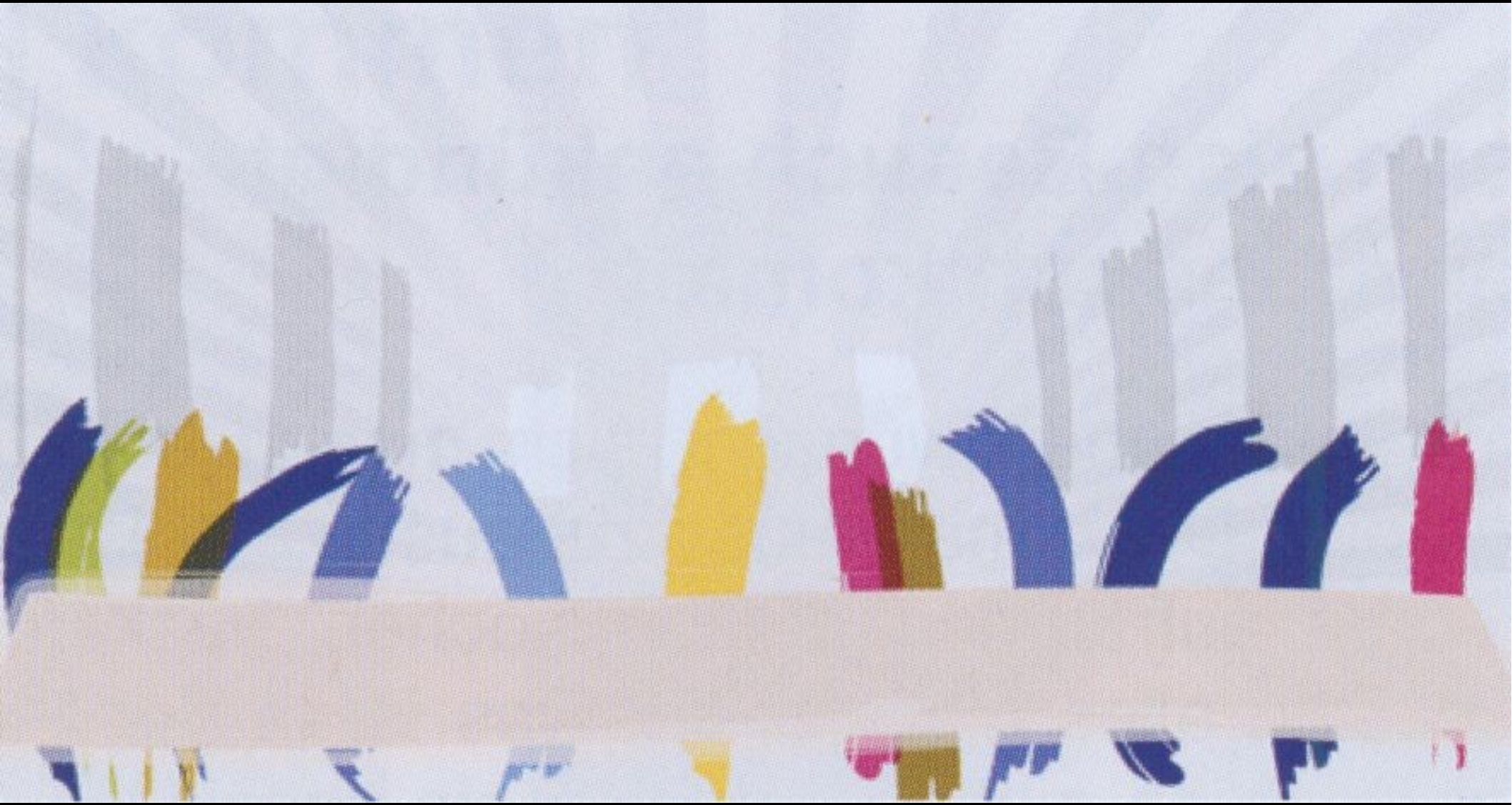
Playing Atari: Results

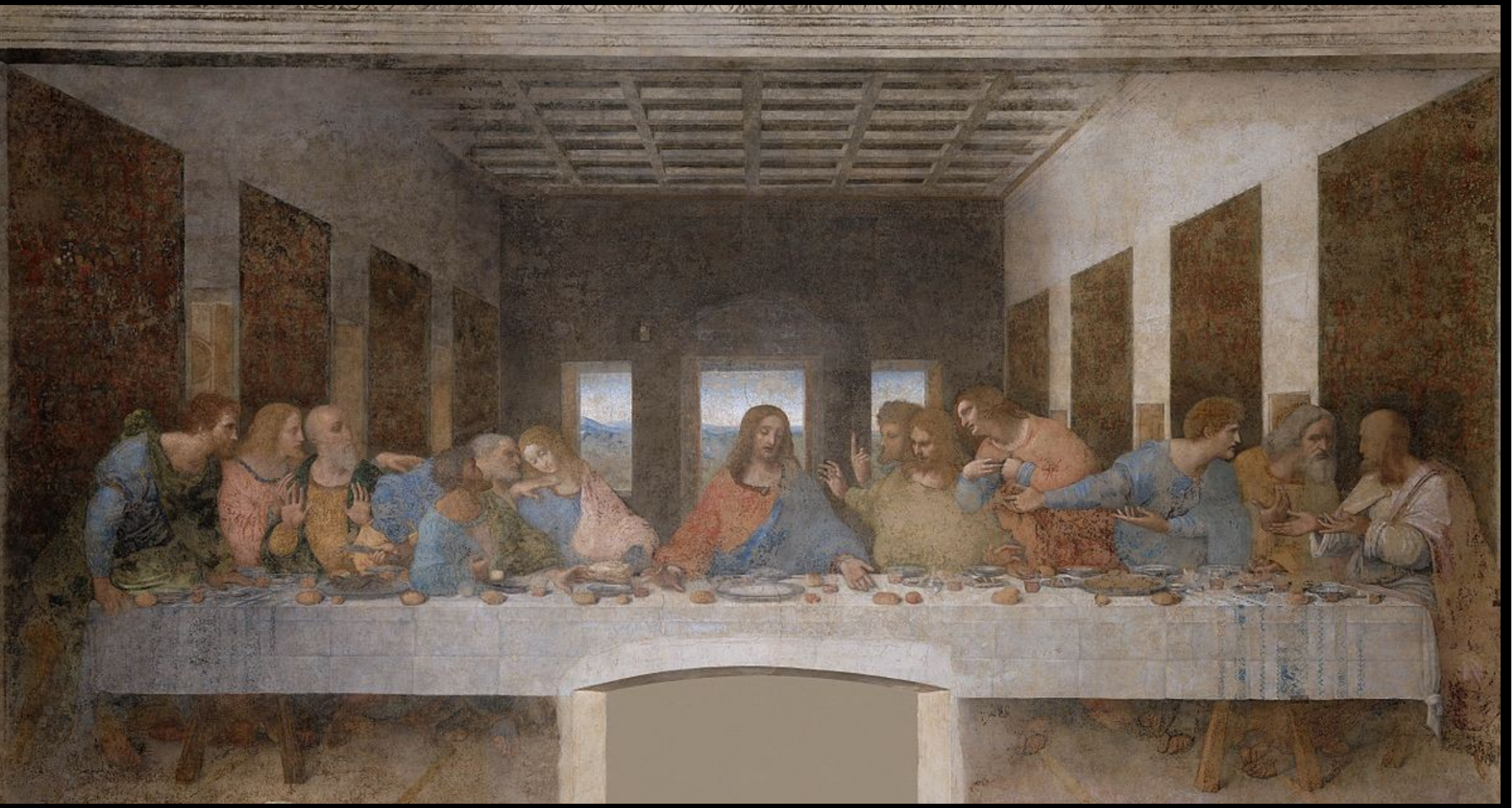
	B. Rider	Breakout	Enduro	Pong	Q*bert	Seaquest	S. Invaders
Random	354	1.2	0	−20.4	157	110	179
Sarsa [3]	996	5.2	129	−19	614	665	271
Contingency [4]	1743	6	159	−17	960	723	268
DQN	**4092**	**168**	**470**	**20**	**1952**	**1705**	**581**
Human	7456	31	368	−3	18900	28010	3690
HNeat Best [8]	3616	52	106	19	1800	920	**1720**
HNeat Pixel [8]	1332	4	91	−16	1325	800	1145
DQN Best	**5184**	**225**	**661**	**21**	**4500**	**1740**	1075

Round-Up

- Neurobiological background
- Script Recognition
- Image Classification
- Strategy Games
- Video Games

What are your problems to be solved?





Leonardo da Vinci, 1494–1498

Marc-Peter Schambach
*1971

Dipl. Physics
PhD Computer Science

Siemens Konstanz, Germany
Postal, Parcel & Airport Logistics
Handwriting Recognition