Visual* Learning of Arithmetic Operations

Yedid Hoshen and Shmuel Peleg

The Hebrew University of Jerusalem

* Visual: Picture In ➡ Picture Out

AAAI 2016

Visual* Learning of Arithmetic Operations

Completely Impractical

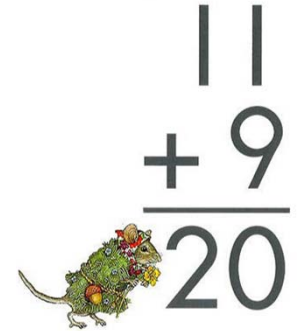## Do we do any Practical Research at The Hebrew University?

Shashua's **MobilEye**, US$10B Company, Autonomous Driving

Peleg's **Briefcam**, doing

Video Synopsis (CVPR 2006)
2:30 min. video clip shown on Taiwan TV 2 weeks ago

Yael Pritch, Alex Rav-Acha

Video Synopsis – Hebrew University

Visual* Learning of Arithmetic Operations

Yedid Hoshen and Shmuel Peleg

The Hebrew University of Jerusalem

* Visual: Picture In ➡ Picture Out

AAAI 2016

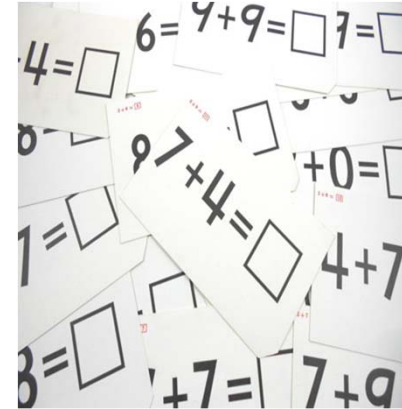# Children Learn Arithmetic Visually

Numbers   Digits   Operations
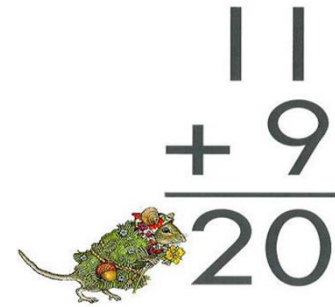
Adults Perform Math Symbolically

- **Concept** of numbers
- **Digits** (0, 1, 5, 9)
- **Operations** (+, -, *, /)

Common Visual Machine Learning
Optical Characcter Recognition

Bitmap

Classifier
$q: X \rightarrow Y$

label

| 1 | 2 | 0 | 7 | 9 | 5 |
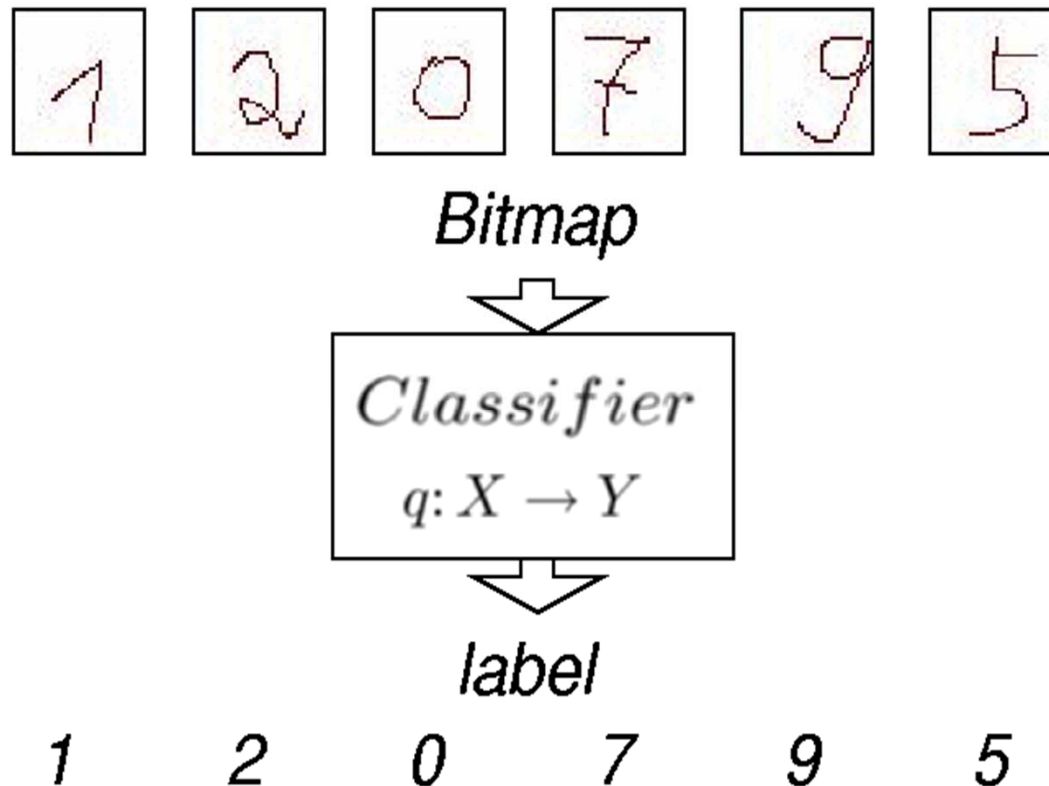
Can AI Methods **Visually** Learn Arithmetic?

Picture In Picture Out

Visually means:

- *Create a picture of N_1*

- *Create a picture of N_2*

- *NN draws picture of $N_1 + N_2$*

No Numbers, Digits, Operations

Traditional Arithmetic
Perception + Cognition + Action

Perception: OCR   9 ⟹ 9   `981925`

Cognition:

981925 **+** 2010445 = 2992370   `2010445`

                      `2992370`

Action: Generate Picture 9 ⟹ 9   `2992370`

Proposed NN: End-to-End Learning of All Tasks
Picture In ➡ Picture Out

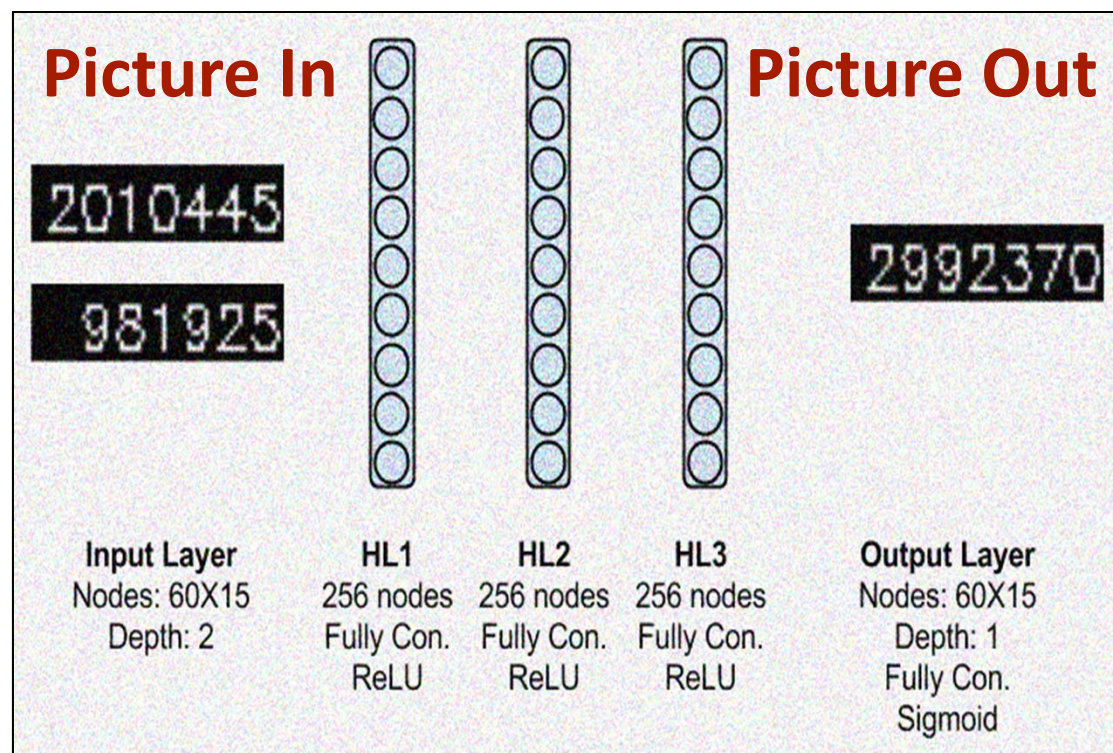# Visual Learning With Neural Networks

Input: Two **pictures**

- Each 60×15

Three Hidden layers:

- Affine 256 nodes
- ReLU

Input Layer	HL1	HL2	HL3	Output Layer
Nodes: 60X15	256 nodes	256 nodes	256 nodes	Nodes: 60X15
Depth: 2	Fully Con.	Fully Con.	Fully Con.	Depth: 1
	ReLU	ReLU	ReLU	Fully Con.
				Sigmoid

Picture In 2010445 981925 — Picture Out 2992370

Output: One **Picture** (60×15), Sigmoid

Train with ADAM, alpha: = 0.1, mom: = 0.9, decay: 1e-4
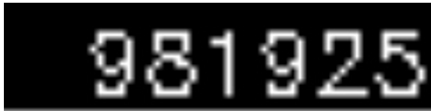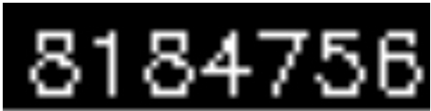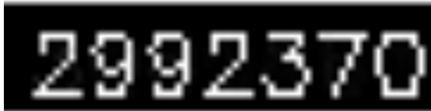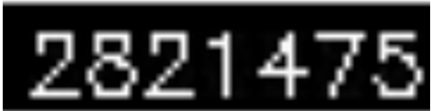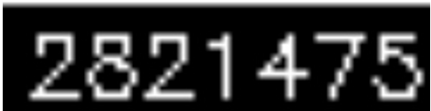
Experiments

- Addition -   range (0 .. 5M)
- Subtraction -   range (0 .. 10M)
- Multiplication - range (0 .. 3.1K)

Training set: 150,000 examples (out of 25×10^{12})

Test set: 30,000 **unseen** examples

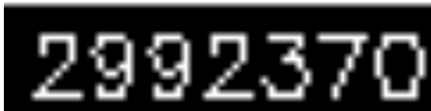# Example of Results

	Addition	Subtraction	Addition + Noise
Picture 1	981925	8184756	4752372
Picture 2	2010445	5363281	230805
NN Result	2992370	2821475	4982677
Correct Result	2992370	2821475	4982677 *

Multiplication Failure Case

Picture 1   2531

Picture 2   1312

NN Result   3368852

Correct Result   3320672

Quantitative Results

Operation	Visual end-to-end		Cognition only: binary (1-hot to 1-hot)	
	No. Layers	%OCR Digit Error	No. Layers	% Error
Addition	3	1.9%	1	1.7%
Subtraction	3	3.2%	1	2.1%
Multiplication	5	71.5%	3	37.6%

- "+" & "-": Both end-to-end & cognition succeeded
- "*"   : Both end-to-end & cognition failed

Implications

- Teach arithmetic to animals
  - Poggio: Very few neurons
  - No language needed

- Explore more operations:
  - Can visual learning be better than symbolic learning?

- Difficulty metric for exams:
  - How hard is a psychometric exam
  - Invariance to language

15