

# Invariantná kategorizácia 2D objektov pomocou Deep Belief Network

Diplomová práca 2011

Bc. Juraj Barič

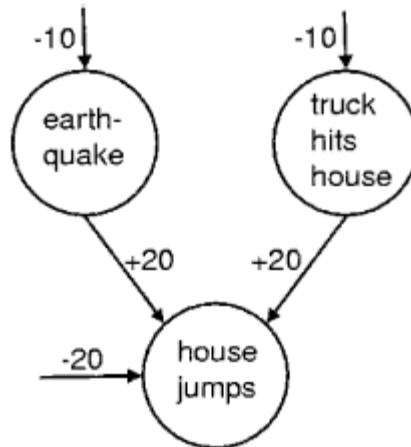
Vedúci dipl. práce doc. Ing. Igor Farkaš, PhD.

# Cieľ práce

- Oboznámiť sa s princípmi hierarchických stochastických učiacich sa systémov so samoorganizáciou, konkrétne s modelom Deep Belief Network
- Implementovať a otestovať DBN na úlohe kategorizácie jednoduchých 2D objektov
- Navrhnuť zlepšenie modelu DBN s cieľom dosiahnuť invariantnosti rozpoznania kategórie objektu (voči posunutiu, rotácii a škále)
- Vyhodnotiť správanie modelu DBN pomocou kvantitatívnych a vizualizačných metód

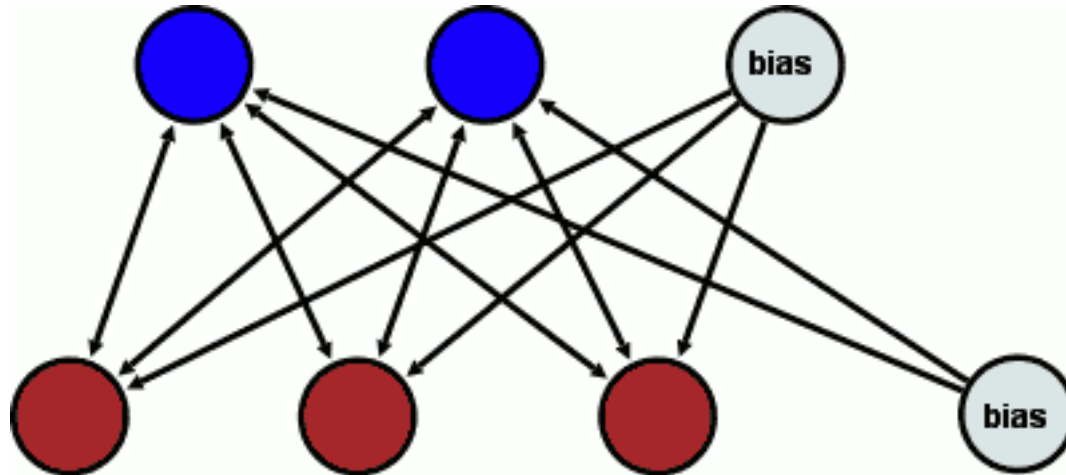
# Motivácia

- Vyhnúť sa „explaining away“ efektu
- Hierarchické učenie
- Unsupervised layer-wise greedy algoritmus – zrýchľuje učenie viacvrstvových štruktúr



# Restricted Boltzmann Machine

- 2-vrstvová symetrická sieť s úplným prepojením (len) medzi vrstvami
- Binárne neuróny
- Generatívny model / Stochaistický model
- učenie so samoorganizáciou



# Restricted Boltzmann Machine

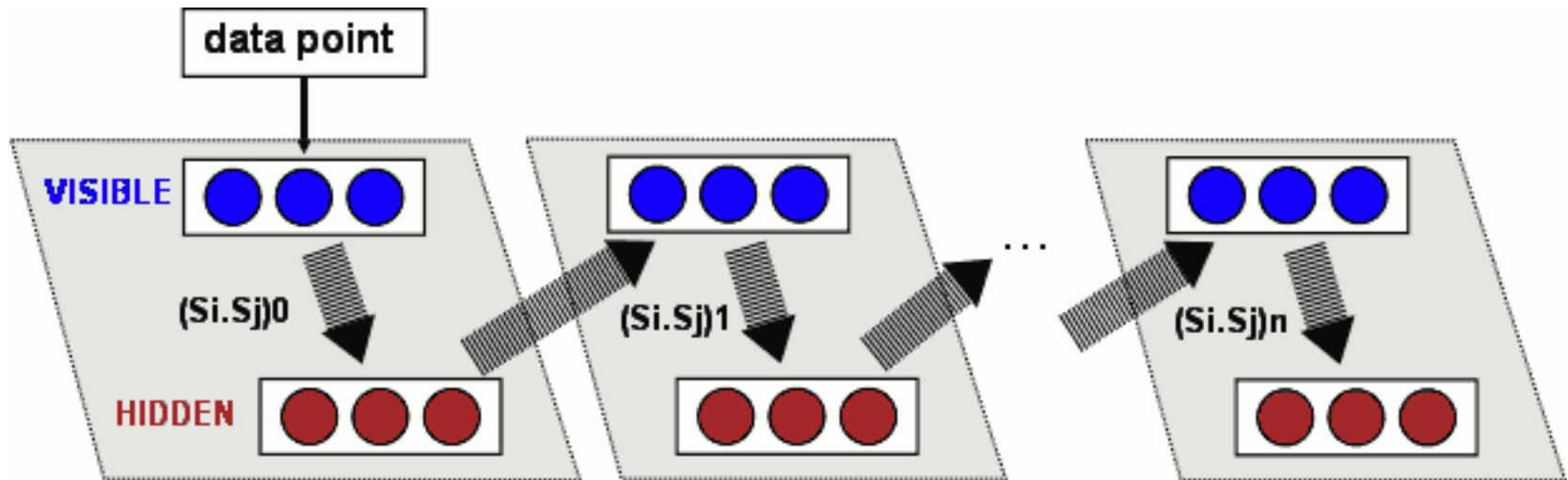
## Notácia

- $W$  – matica váh
- $b$  – bias pre skrytú vrstvu (indexovaný  $j$ )
- $c$  – bias pre viditeľnú vrstvu (indexovaný  $i$ )
- $v$  – viditeľná vrstva
- $h$  – skrytá vrstva

# Restricted Boltzmann Machine

- Sigmoidná funkcia  $\sigma(x) = 1 / (1 + \exp(-x))$
- $P(h_j == 1) = \sigma( (v^*W)_j + b_j )$
- $P(v_i == 1) = \sigma( (W^*h)_i + c_i )$
- $E(v, h) = - v^*W^*h - b^*v - c^*h$
- $P(v, h) \approx \exp(-E(v, h))$
- $P(v) = \sum P(v, h)$

# RBM - učenie



$$\Delta w_{ij} = \varepsilon ( \langle s_i s_j \rangle^0 - \langle s_i s_j \rangle^\infty )$$

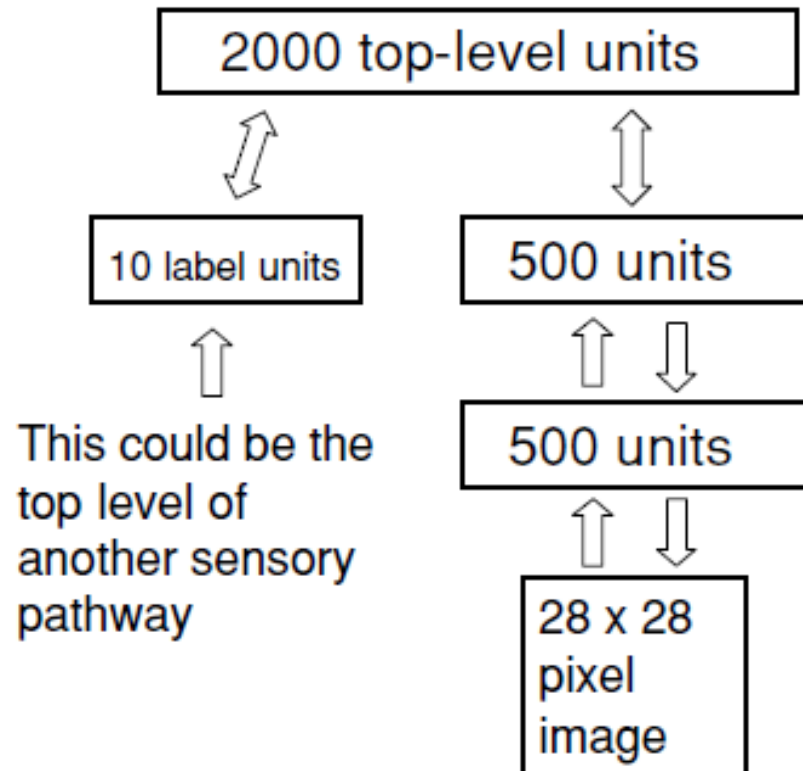
$$\Delta w_{ij} \propto s_j (s_i - \hat{s}_i)$$

# Deep Belief Network

- viacvrstvová sieť (pozostáva z viacerých RBM)
- Stochastický model
- Symetrické váhy
- Generatívny model



# Deep Belief Network



# Predošlé testovanie DBN

- Facial expressions generating
- Spam recognition
- MNIST digits recognition
- Iné menšie úlohy

# Predošlé výsledky DBN (spam filtering)

- Miera úspešnosti:

$$Acc = \frac{|S \rightarrow S| + |L \rightarrow L|}{|S \rightarrow S| + |L \rightarrow L| + |S \rightarrow L| + |L \rightarrow S|}$$

$$HR = \frac{|L \rightarrow L|}{|L \rightarrow L| + |L \rightarrow S|} \quad (8) \quad SR = \frac{|S \rightarrow S|}{|S \rightarrow S| + |S \rightarrow L|}$$

$$HP = \frac{|L \rightarrow L|}{|L \rightarrow L| + |S \rightarrow L|} \quad (10) \quad SP = \frac{|S \rightarrow S|}{|S \rightarrow S| + |L \rightarrow S|}$$

- LingSpam (481 / 2893 - spam ratio: 16.6%)
- SpamAssassin (1897 / 6047 - spam ratio: 31.3%)
- EnronSpam (1500 / 5172 - spam ratio: 29.0%)

# Predošlé výsledky DBN

Performance Measure	LingSpam	
	DBN 1500-50-50-200-2	SVM C = 1
Accuracy	<b>99.45%</b>	99.24%
Spam Recall	<b>98.54%</b>	96.67%
Spam Precision	98.2%	<b>98.74%</b>
Ham Recall	99.63%	<b>99.75%</b>
Ham Precision	<b>99.71%</b>	99.35%

Table 2. DBNs vs. SVMs on SpamAssassin.

Performance Measure	SpamAssassin	
	DBN 1000-50-50-200-2	SVM C = 10
Accuracy	<b>97.5%</b>	97.32%
Spam Recall	<b>95.51%</b>	95.24%
Spam Precision	<b>96.4%</b>	96.14%
Ham Recall	<b>98.39%</b>	98.24%
Ham Precision	<b>98.02%</b>	97.89%

Table 3. DBNs vs. SVMs on Enron1.

Performance Measure	Enron1	
	DBN 1000-50-50-200-2	SVM C = 1
Accuracy	<b>97.43%</b>	96.92%
Spam Recall	96.47%	<b>97.27%</b>
Spam Precision	<b>94.94%</b>	92.74%
Ham Recall	<b>97.83%</b>	96.78%
Ham Precision	98.53%	<b>98.84%</b>

# Predošlé výsledky DBN (MNIST recognition)

Version of MNIST Task	Learning Algorithm	Test Error %
Permutation invariant	Our generative model: 784 → 500 → 500 ↔ 2000 ↔ 10	1.25
Permutation invariant	Support vector machine: degree 9 polynomial kernel	1.4
Permutation invariant	Backprop: 784 → 500 → 300 → 10 cross-entropy and weight-decay	1.51
Permutation invariant	Backprop: 784 → 800 → 10 cross-entropy and early stopping	1.53
Permutation invariant	Backprop: 784 → 500 → 150 → 10 squared error and on-line updates	2.95
Permutation invariant	Nearest neighbor: all 60,000 examples and L3 norm	2.8
Permutation invariant	Nearest neighbor: all 60,000 examples and L2 norm	3.1
Permutation invariant	Nearest neighbor: 20,000 examples and L3 norm	4.0
Permutation invariant	Nearest neighbor: 20,000 examples and L2 norm	4.4
Unpermuted images; extra data from elastic deformations	Backprop: cross-entropy and early-stopping convolutional neural net	0.4
Unpermuted de-skewed images; extra data from 2 pixel translations	Virtual SVM: degree 9 polynomial kernel	0.56
Unpermuted images	Shape-context features: hand-coded matching	0.63
Unpermuted images; extra data from affine transformations	Backprop in LeNet5: convolutional neural net	0.8
Unpermuted images	Backprop in LeNet5: convolutional neural net	0.95

# Predošlé výsledky DBN

Dataset	SVM <sub>rbf</sub>	SVM <sub>poly</sub>	NNet	DBN-1	SAA-3	DBN-3
<i>mnist-basic</i>	<b>3.03±0.15</b>	3.69±0.17	4.69±0.19	3.94±0.17	3.46±0.16	<b>3.11±0.15</b>
<i>mnist-rot</i>	<b>10.38±0.27</b>	13.61±0.30	17.62±0.33	12.11±0.29	11.43±0.28	12.30±0.29
<i>mnist-back-rand</i>	14.58±0.31	16.62±0.33	20.04±0.35	9.80±0.26	11.28±0.28	<b>6.73±0.22</b>
<i>mnist-back-image</i>	22.61±0.37	24.01±0.37	27.41±0.39	<b>16.15±0.32</b>	23.00±0.37	<b>16.31±0.32</b>
<i>mnist-rot-back-image</i>	32.62±0.41	37.59±0.42	42.17±0.43	31.84±0.41	<b>24.09±0.37</b>	28.51±0.40
<i>rectangles</i>	<b>2.15±0.13</b>	<b>2.15±0.13</b>	7.16±0.23	4.71±0.19	<b>2.41±0.13</b>	2.60±0.14
<i>rectangles-image</i>	24.04±0.37	24.05±0.37	33.20±0.41	23.69±0.37	24.05±0.37	<b>22.50±0.37</b>
<i>convex</i>	19.13±0.34	19.82±0.35	32.25±0.41	19.92±0.35	<b>18.41±0.34</b>	<b>18.63±0.34</b>

# Diplomová práca

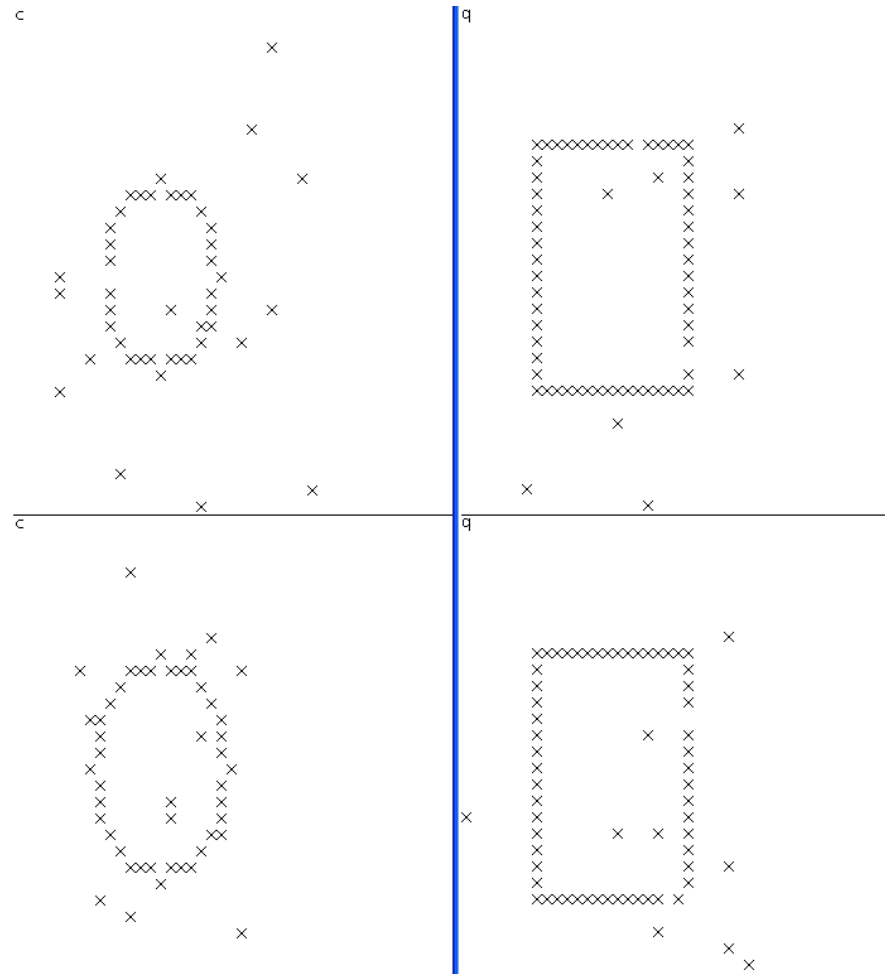
- Matrix tester – porovnanie performance násobení matic
- Generátor tréningových dát
- DBN tester
- WeightsProfileViewer

# Výsledky DBN

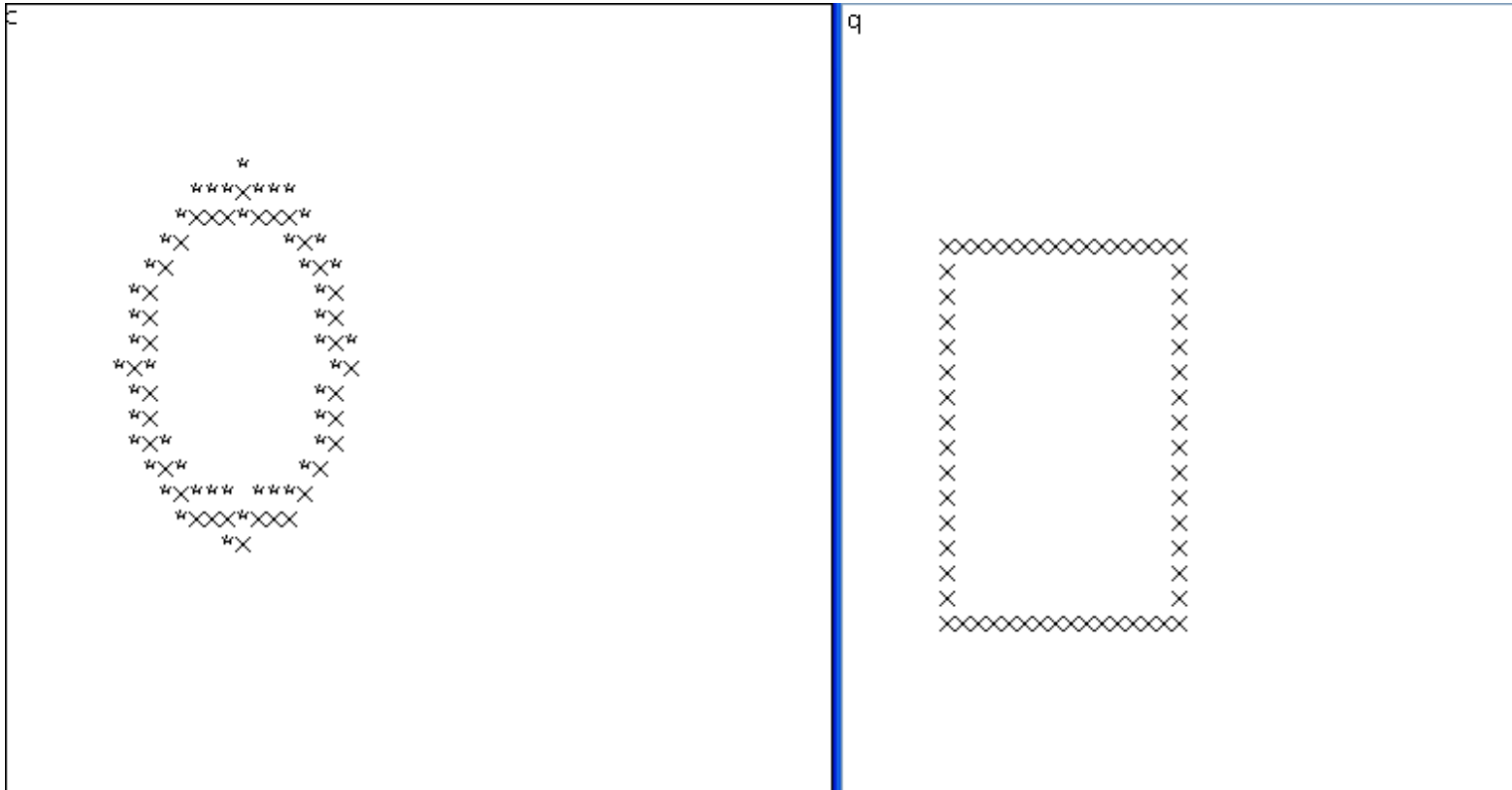
- Kategorizácia 3 druhov obrázkov (štvorec, trojuholník, kruh) pri šume do 10 %
- Kategorizácia pri zvýšenom šume
- Kategorizácia pozične variabilných obrázkov
- Kategorizácia uhlovo variabilných obrázkov



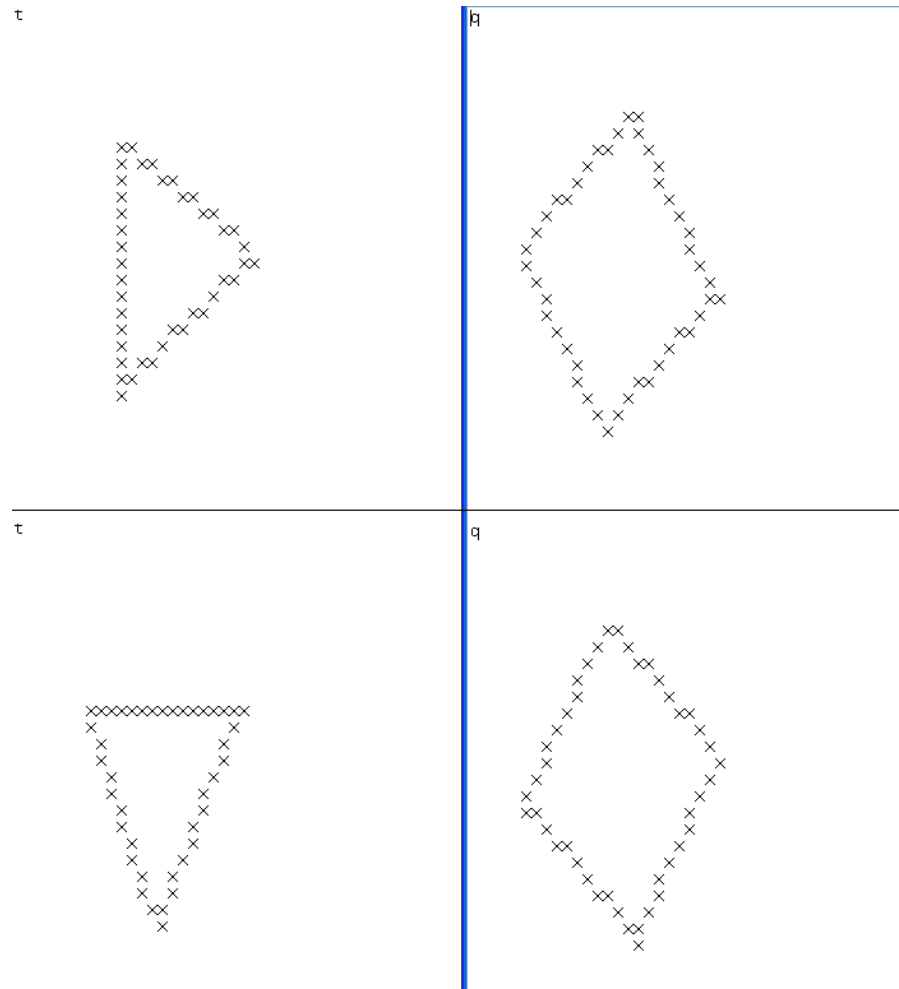
# Príklady tréningových dát



# Príklady trérovacích dát



# Príklady tréningových dát



# Výsledky – pri šume do 10 %

- Počet výstupov : 3
- Alfa : 0.7
- Veľkosť obrázkov : 30 x 30
- Počet skrytých vrstiev : 3
- Veľkosti skrytých vrstiev : 500 x 500 x 1500
- Počet iterácií : 30

## Počet chýb počas testu

- krúžky: 0/50
- štvorce: 0/50
- trojuholníky: 0/50

# Výsledky – pozičná variabilita úplná

- Počet výstupov : 3
- Alfa : 0.7
- Veľkosť obrázkov : 20 x 20
- Počet skrytých vrstiev : 3
- Veľkosti skrytých vrstiev : 500 x 500 x 1500
- Počet iterácií : 30

## Počet chýb počas testu

- krúžky: 1/20
- štvorce: 20/20
- trojuholníky: 11/20

# Výsledky – pozičná variabilita úplná

- Počet výstupov : 3
- Alfa : 0.3
- Veľkosť obrázkov : 20 x 20
- Počet skrytých vrstiev : 3
- Veľkosti skrytých vrstiev : 500 x 500 x 1500
- Počet iterácií : 60
- Počet tréningových dát : 75
- Tréningové dáta permutované

Počet chýb počas testu

- krúžky: 8/20
- štvorce: 19/20
- trojuholníky: 13/20

# Výsledky – pozičná variabilita čiastočná (fuzzy hrany objektov)

- Čas výpočtu v sekundách : 745.344
- Počet výstupov : 3
- Alfa : 0.15
- Veľkosť obrázkov : 30 x 30
- Počet skrytých vrstiev : 3
- Veľkosti skrytých vrstiev : 700 x 700 x 1200
- Počet iterácií : 200
- Počet tréningových dát : 60
  
- Počet chýb počas testu : 0
- Celkový počet testov : 15

# Výsledky – uhlová variabilita úplná

- Čas výpočtu v sekundách : 26.718
- Počet výstupov : 3
- Alfa : 0.15
- Veľkosť obrázkov : 30 x 30
- Počet skrytých vrstiev : 3
- Veľkosti skrytých vrstiev : 200 x 200 x 600
- Počet iterácií : 100
- Počet tréningových dát : 30
  
- Počet chýb počas testu : 0
- Celkový počet testov : 10



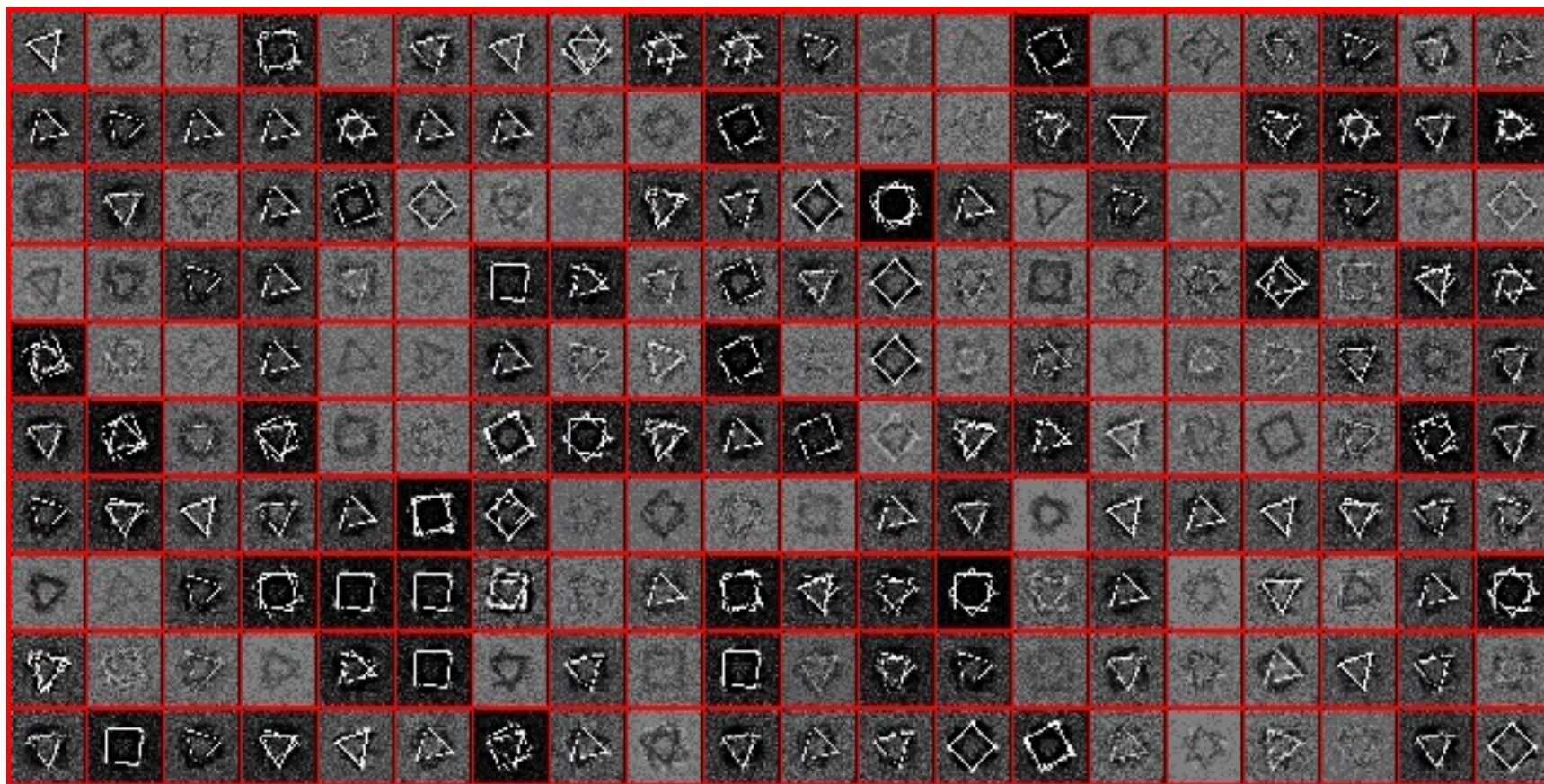
# Výsledky – uhlová variabilita úplná

- Čas výpočtu v sekundách : 8.25
- Počet výstupov : 3
- Alfa : 0.15
- Veľkosť obrázkov : 30 x 30
- Počet skrytých vrstiev : 3
- Veľkosti skrytých vrstiev : 200 x 200 x 600
- Počet iterácií : 100
- Počet tréningových dát : 10
  
- Počet chýb počas testu : 10
- Celkový počet testov : 30

# Váhový profil DBN – uhlová variabilita úplná

- Čas výpočtu v sekundách : 13.797
- Počet výstupov : 3
- Alfa : 0.15
- Veľkosť obrázkov : 30 x 30
- Počet skrytých vrstiev : 3
- Veľkosti skrytých vrstiev : 200 x 200 x 600
- Počet iterácií : 100
- Počet tréningových dát : 16
  
- Počet chýb počas testu : 3
- Celkový počet testov : 24

# Váhový profil DBN – uhlová variabilita úplná

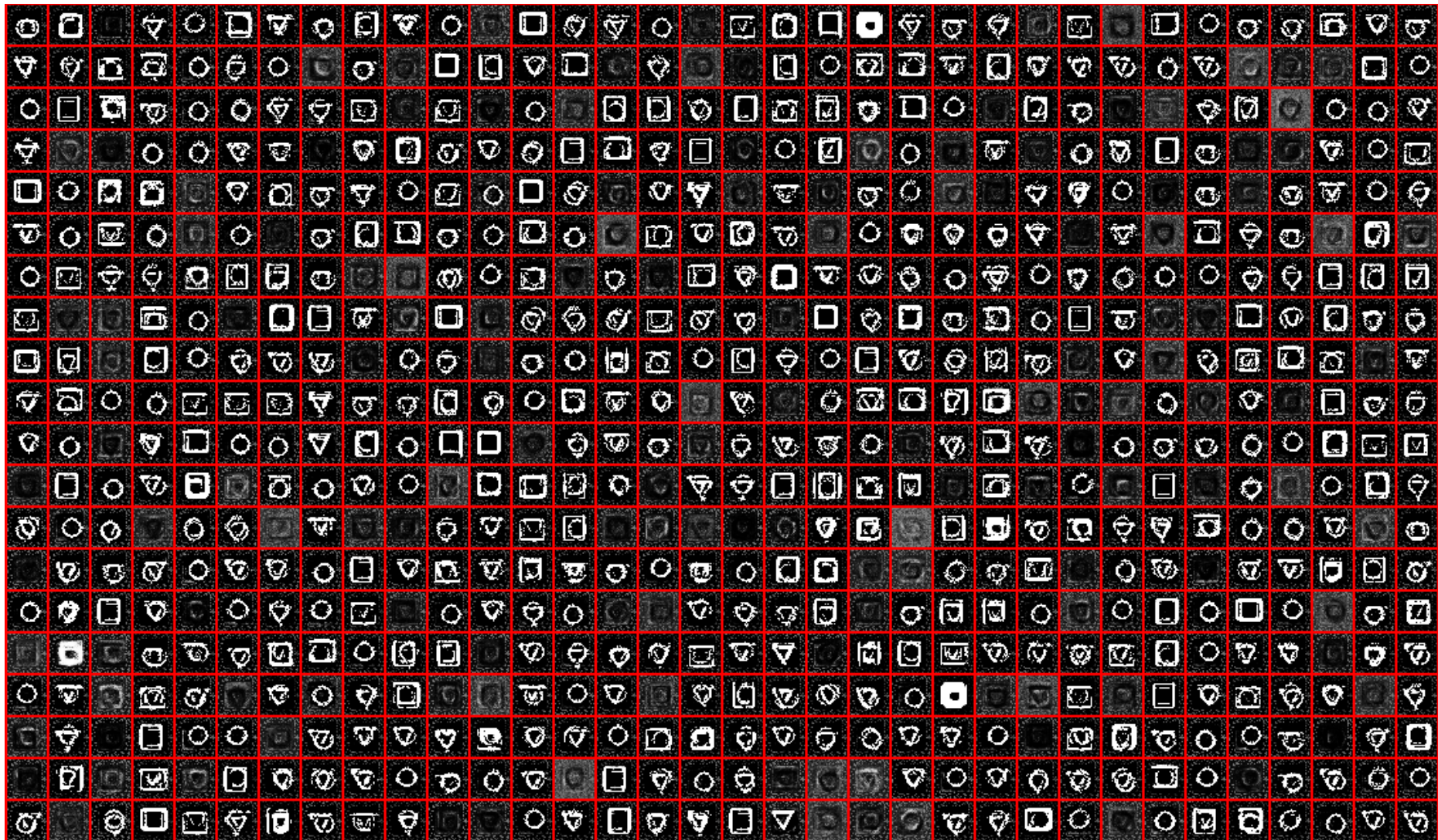


# Váhový profil DBN – pozičná variabilita čiasočná (fuzzy hrany)

- Čas výpočtu v sekundách : 745.344
- Počet výstupov : 3
- Alfa : 0.15
- Veľkosť obrázkov : 30 x 30
- Počet skrytých vrstiev : 3
- Veľkosti skrytých vrstiev : 700 x 700 x 1200
- Počet iterácií : 200
- Počet tréningových dát : 60
  
- Počet chýb počas testu : 0
- Celkový počet testov : 15



# Váhový profil DBN – pozičná variabilita čiasočná (fuzzy hrany)

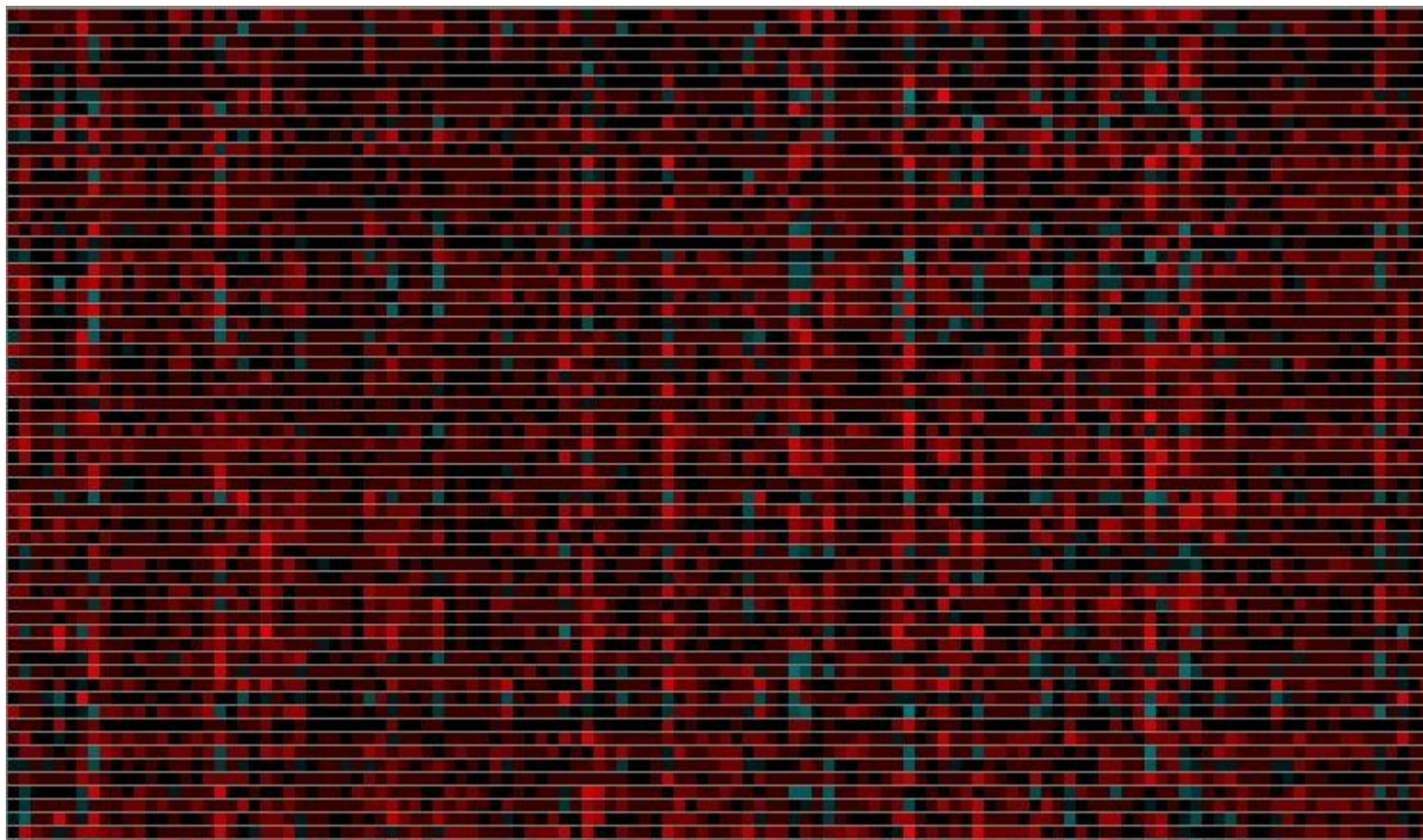


# Rekonstrukcia

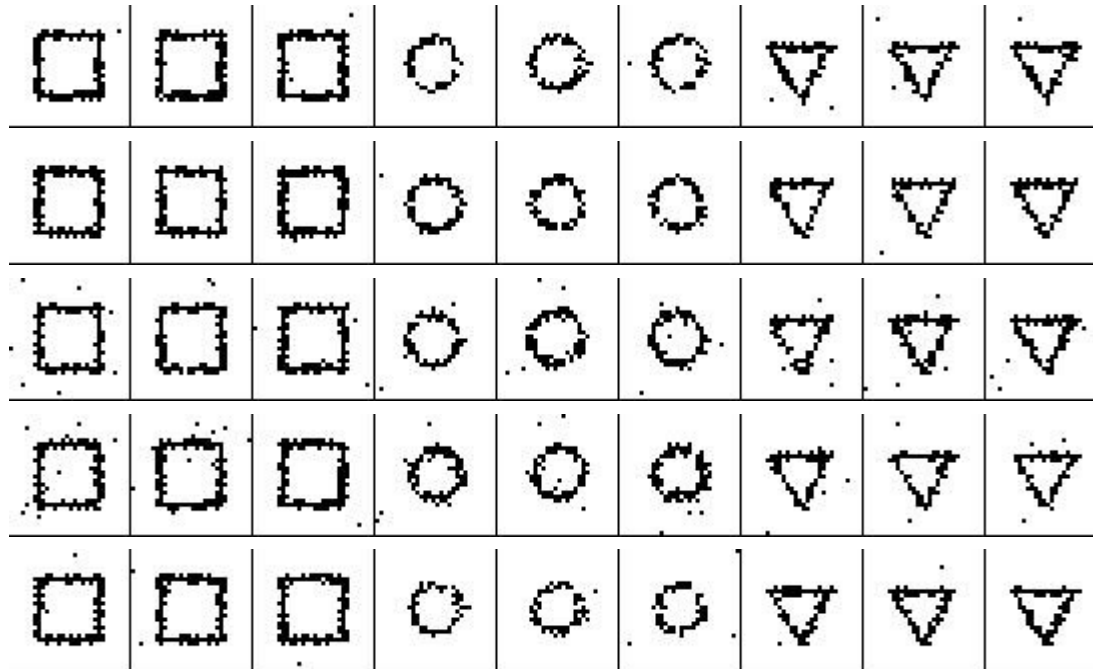
- Problém pri jedinom zapnutom neuróne
- Veľa záporných váh
- Riešenie: zosilnenie signálu z výstupu



# Rekonstrukcia

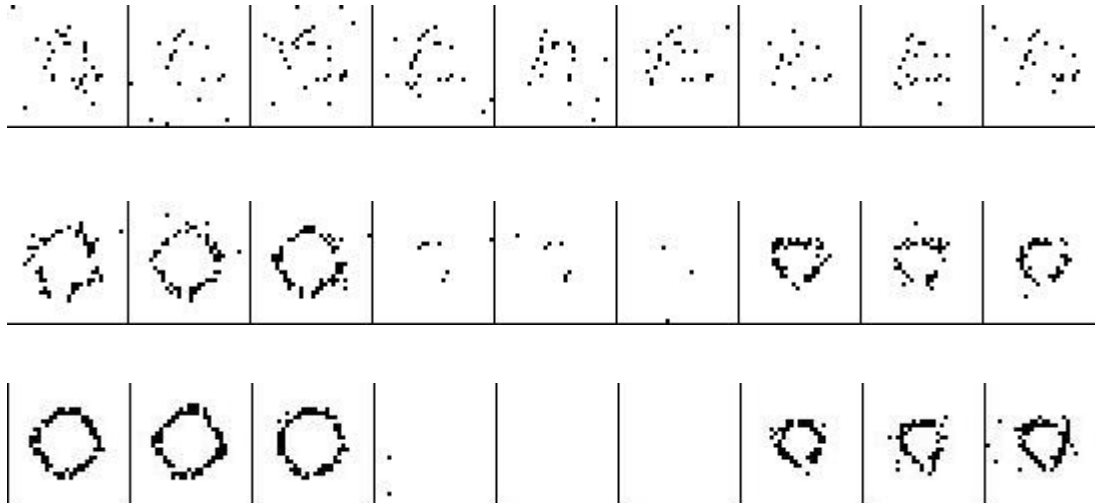


# Rekonstrukcja



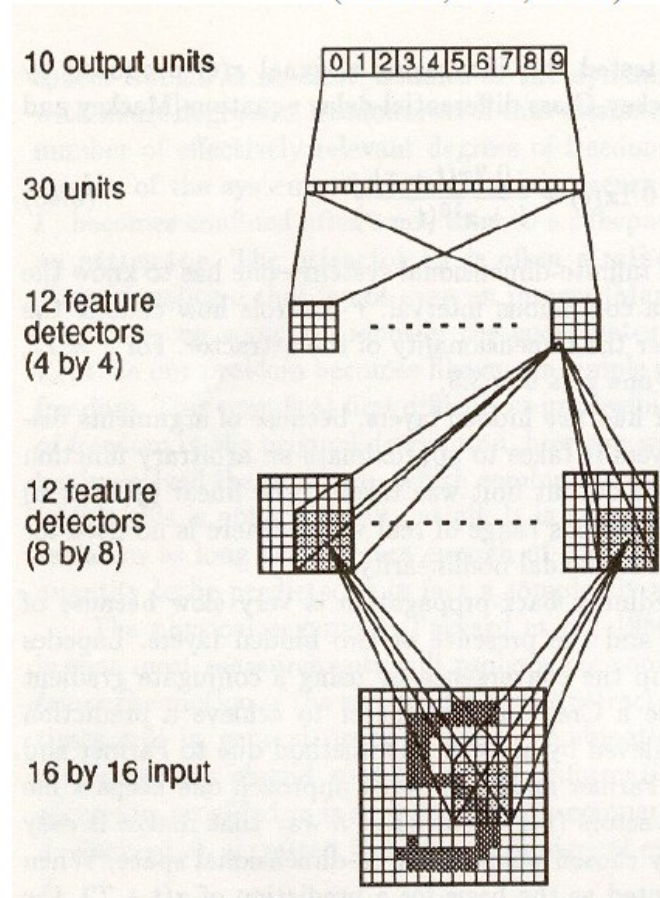


# Rekonstrukcja



# DBN so sdiel'anými váhami a receptívnymi poliami

(LeCun, et al, 1989)



# DBN so sdiel'anými váhami a receptívnymi poliami

- Inšpirácia modelom LeCun et al. (1989)
- Sdiel'anie váh (weight-sharing) redukuje počet parametrov modelu, podporuje generalizáciu
- DBN ignoruje 2D štruktúru
- Problém: Gibbs sampling spôsobí, že vstupný neurón ovplyvní aj ten skrytý neurón, s ktorým nie je spojený?

# Použité materiály

1. Hinton : A Fast Learning Algorithm for Deep Belief Nets, 2006
2. Grigorios Tzortzis and Aristidis Likas : Deep Belief Networks for Spam Filtering
3. Bengio : An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variations
4. Hinton : Generating facial expressions with DBN
5. Le cun : Handwritten digit recognition with a back-propagation Network, 1990