Deep Spiking Neural Network for Visual Pattern Recognition

Daqi Liu

School of Computer Science, University of Lincoln

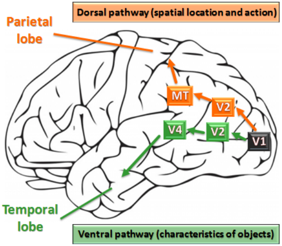
October 2, 2017

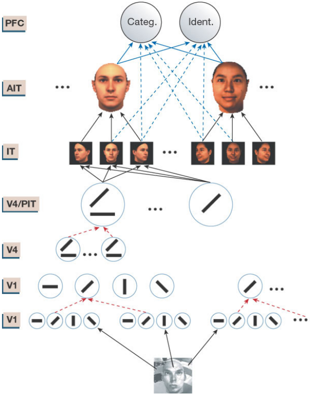
Outline

1. Introduction
2. Main Works
3. Future Works

Visual Pathway

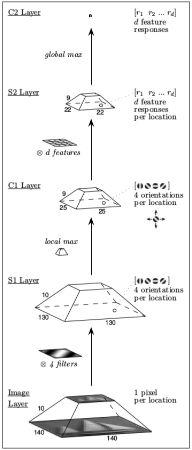
- Ventral stream





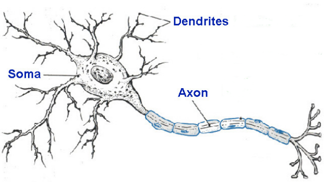
Deep Learning

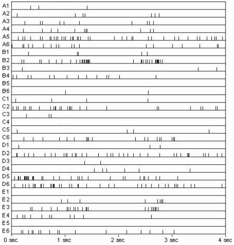
- Convolutional neural network (HMAX model)



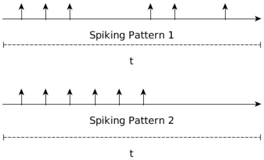
Spiking Neural Network

- Spiking neurons



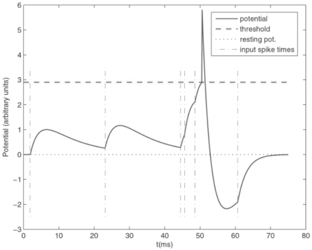


- ANN vs SNN



Neuron Model

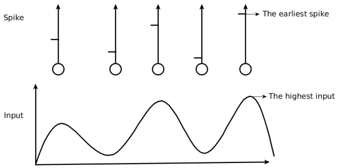
- Leaky integrate-and-fire (LIF) model



- Define neuron behaviors
- Coincidence detector

Spiking Coding Scheme

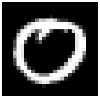
- Spiking rate vs spiking timing sequence
- Rank order coding (ROC)

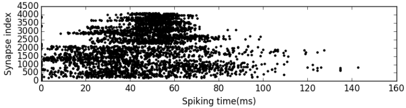


- Neuron is only allowed to fire at most once
- First spike wave is enough for further processing

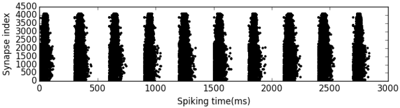
Spiking Coding Scheme

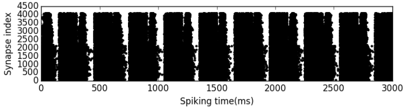
- One input image and its spiking pattern





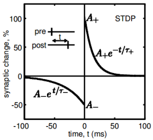
- Spiking pattern sequence

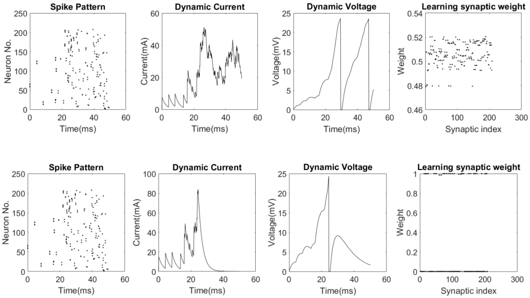




Learning Method

- Spike-timing dependent plasticity (STDP)



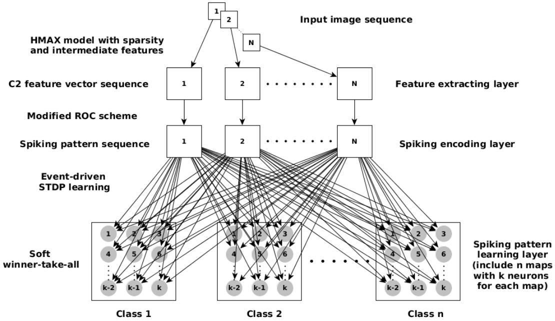


Event-driven Continuous STDP Learning (ECS)

- **State-of-the-art Methods**
- **Spiking rate-based models**
- Vanishing/exploding gradient problem
- Over-fitting, not robust
- Incorporate global error information
- Require long processing time
- Not biologically plausible
- **Spiking timing-based models**
- Require supervisory signal, no strong experimental confirmation
- STDP is used as a local feature extractor
- Not biologically plausible

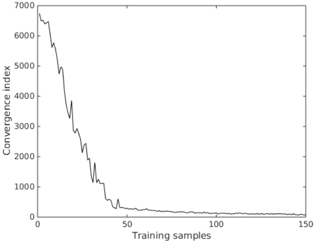
Event-driven Continuous STDP Learning (ECS)

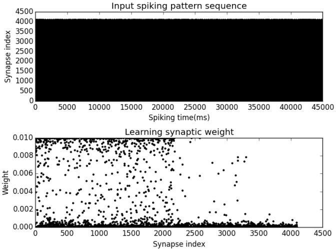
- ECS architecture



Event-driven Continuous STDP Learning (ECS)

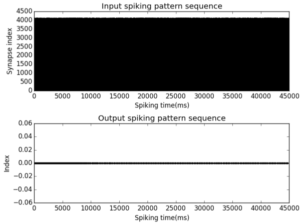
- Convergence analysis

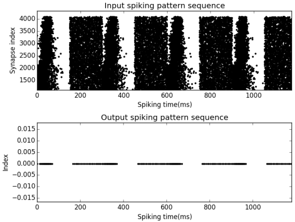


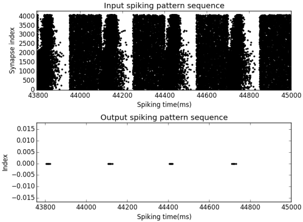


Event-driven Continuous STDP Learning (ECS)

- Robustness analysis







Event-driven Continuous STDP Learning (ECS)

- Experimental results

TABLE IV: Classification accuracy performance using different methods on MNIST database.

Spiking Coding-type	Architecture	Preprocessing	(Un-)supervised	Learning Rule	Performance: Simple random sampling [a]	Performance: Exhaustive [b]
Time-based	Spiking convolutional neural network	Modified HMAX	Supervised	**ECS(this paper)**	**89%**	**93.0%**
	Two layer network[10]	Simplified HMAX	Supervised	Tempotron rule	79.0%	N/A
	Two layer network[11]	Simplified HMAX	Supervised	Tempotron rule	N/A	91.3%
Rate-based	Dendritic neurons[4]	Thresholding	Supervised	Morphology learning	N/A	90.3% [d]
	Spiking RBM[5]	None	Supervised	Contrastive divergence, linear classifier	N/A	89.0%
	Spiking RBM[6]	Enhanced training set to 120,000 examples	Supervised	Contrastive divergence	N/A	89.0%
	Spiking convolutional neural network[7]	None	Supervised	Backpropagation	N/A	99.1%
	Spiking RBM[8]	Thresholding	Supervised	Contrastive divergence	N/A	92.6% [c]
	Spiking RBM[8]	Thresholding	Supervised	Contrastive divergence	N/A	91.9% [c]
	Two layer network[9]	Edge-detection	Supervised	STDP with calcium variable	N/A	96.5% [e]
	Multi-layer hierarchical neural network[1]	Orientation-detection	Supervised	STDP with calcium variable	N/A	91.6%
	Two layer network[2]	None	Unsupervised	Rectangular STDP	N/A	93.5%
	Two layer network[3]	None	Unsupervised	Exponential STDP	N/A	95.0%

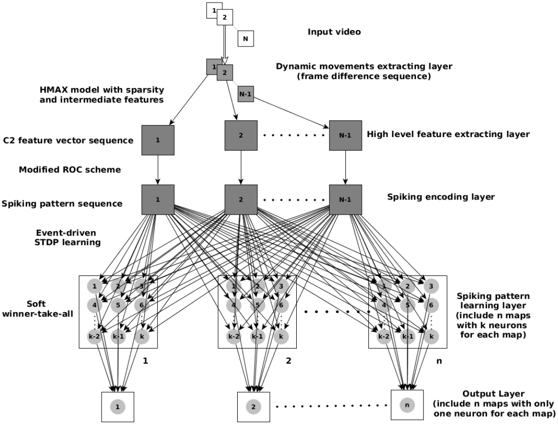
[a] Simple random sampling performance has been generated by averaging 10 random tests using 50 random training samples per class and 100 random testing samples, which is suitable for real-time learning since the whole database is impossible to obtain in most real scenarios.
[b] Exhaustive performance shows the ideal experimental results by using whole 60000 training samples and 10000 testing samples within MNIST database.
[c] The authors only use 1000 testing samples to obtain the performance
[d] The authors only use 5000 testing samples to obtain the performance
[e] The authors use 10000 randomly chosen samples from MNIST database instead of the dedicated testing database

Video-based Disguise Face Recognition (VDFR)

- **State-of-the-art VFR Methods**
- Set-based methods
- Sequence-based methods
- **Research Problems**
- It is often hard to obtain the ideal face frames
- Rely on the features which will be difficult to capture when there are invisible areas
- Does not incorporate disguise variations in current databases

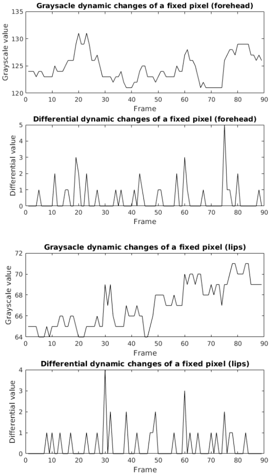
Video-based Disguise Face Recognition (VDFR)

- VDFR architecture



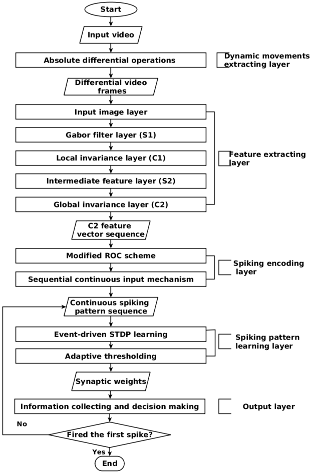
Video-based Disguise Face Recognition (VDFR)

- dynamic facial movements



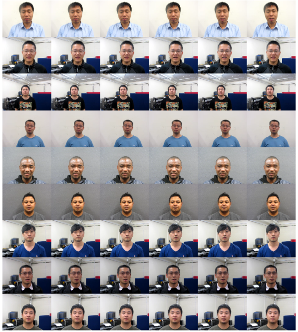
Video-based Disguise Face Recognition (VDFR)

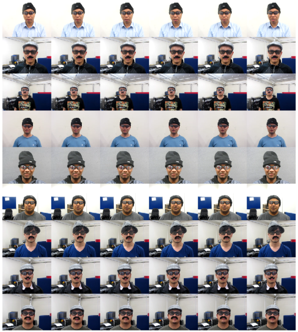
- Flowchart



Video-based Disguise Face Recognition (VDFR)

- MakeFace database





Video-based Disguise Face Recognition (VDFR)

- Experimental results

TABLE III: Correct classification performance using different number of training and testing video clips (%).

Number of training video clips	Number of testing video clips	Performance
1	4	92.5
2	3	100
3	2	100

TABLE IV: Classification performances of two different methods on testing video clips with disguise (%).

Method	Correct rate	Wrong rate	Unknown rate
CNN [29]	93.1 ± 1.35	6.9 ± 1.35	0
Proposed VDFR method	95.2 ± 2.65	4.8 ± 2.65	0

• Note: The classification rate has been computed by averaging 10 random tests. Furthermore, we have conducted a Wilcoxon signed-rank test on the correct classification performances by using the above two methods and computed the significance level $p-value$ (0.03429). Such significance level ($p-value < 0.05$) indicates that the two correct classification performances are statistically different.

TABLE V: Classification performances of two different methods on testing mixed video clips (%).

Method	Correct rate	Wrong rate	Unknown rate
CNN [29]	96.7 ± 0	3.3 ± 0	0
Proposed VDFR method	100 ± 0	0 ± 0	0

• Note: The classification rate has been computed by averaging 10 random tests.

TABLE VI: Classification performances of two different methods on testing unknown video clips (%).

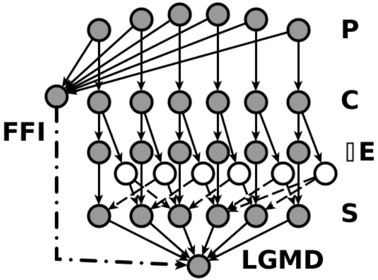
Method	Correct rate	Wrong rate	Unknown rate
CNN [29]	0	0	100 ± 0
Proposed VDFR method	0	0	100 ± 0

• Note: The classification rate has been computed by averaging 10 random tests.

A Spiking LGMD Model for Collision Detection

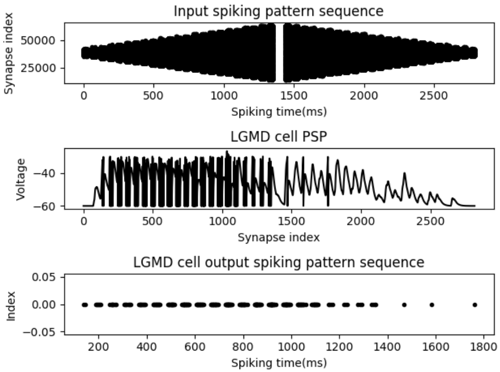
- **Research Problems of Current Models**
  - Only incorporate spiking concept in final decision making step
  - Do not incorporate spiking neural network during detection
  - Do not generate the collision selection observed in LGMD cell
- **Proposed model**
  - Add a spiking encoding layer behind the P layer
  - Incorporate a Poisson point process to generate spike trains
  - Spikes are the only accepted information medium
  - Use an exponential level conductance-based LIF model within S layer

A Spiking LGMD Model for Collision Detection

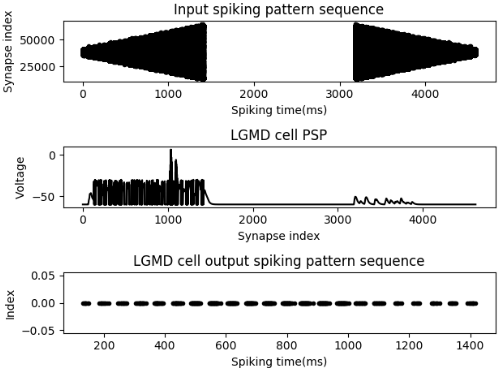


- Differentiate the post-synaptic membrane potentials generated when approaching and receding the object
- Generate a similar collision selection as the real LGMD cell
- Compare with current models, it is biologically plausible

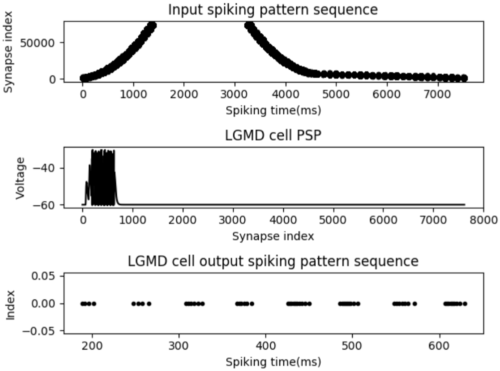
A Spiking LGMD Model for Collision Detection



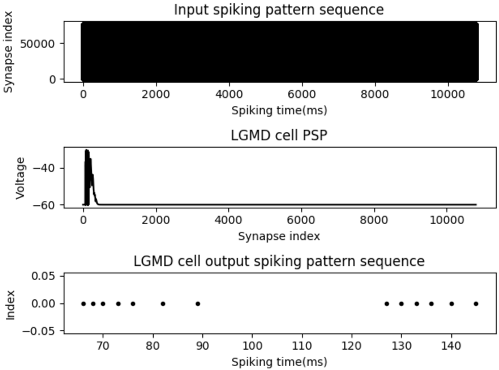
A Spiking LGMD Model for Collision Detection



A Spiking LGMD Model for Collision Detection



A Spiking LGMD Model for Collision Detection



Future Works

- Finish the spiking LGMD model for collision detection
- Investigate the proposed VDFR method against a complex moving background
- Propose an alternative competitive learning method to replace the current STDP learning rule

Thank you!