Course: Multi-modal Data Science and Engineering (MDSE)

Course organiser: Prof. Shihua Zhou

Course presenter

Prof Nikola Kasabov

Visiting Professor at Dalian University

Life FIEEE, FRSNZ, FINNS, DVF RAE UK

Founding Director KEDRI

Professor, Auckland University of Technology, NZ

George Moore Chair/Professor, Ulster University, UK

Honorary Professor, University of Auckland NZ , Peking University China

Visiting Professor IICT/Bulgarian Academy of Sciences and Teesside University UK

Doctor Honoris Causa Obuda University Budapest

Director, Knowledge Engineering Consulting Ltd (https://www.knowledgeengineering.ai)

Assistant

Doctorant Iman AbouHassan

iabouhassan@tu-sofia.bg

abouhassan.iman@gmail.com

Multimodal Data Science and Engineering (MDSE)

Course by research papers.
Every topic will include:

1. Topic/task/problem specification
2. Previously published methods for solving the problem
3. Description of the new method and the publication where it is published
4. Software implementation, experimental results and discoveries
5. Applications
6. Future work to be done for this problem and questions for individual work

Additional materials: - relevant papers;
- https://www.knowledgeengineering.ai/china

ZOOM link for all lectures: https://us05web.zoom.us/j/4658730662?pwd=eFN0eHRCN3o4K0FaZ0lqQmN1UUgydz09

Full List of Topics/Lectures:

1. Introduction to the course: What is MDSE and why we need it?

2. Methods for MDSE:

- *paper: S.Budhraja, B.Singh, S.Tan, M.Dobrojeh, Z.Doborjeh, W.Goh, E.Lai and N.Kasabov, Mosaic LSM: A Liquid State Machine Approach for Multimodal Longitudinal Data Analysis, Proc. International Joint Conference on Neural Networks (IJCNN), Gold Coast, Australia, 2023, pp. 1-8, doi: https://doi.org/10.1109/IJCNN54540.2023.10191256; https://ieeexplore.ieee.org/document/10191256. IEEE, 2023, ISBN:978-166548867-9*
- *Software NeuGems: https://kedri.aut.ac.nz/news-and-events/introducing-neurogems*

3. MDSE for biomedical engineering

- *Paper 1: M. Doborjeh, N. Kasabov, Z. Doborjeh, R. Enayatollahi, E. Tu, A. H. Gandomi, Personalised modelling with spiking neural networks integrating temporal and static information, Neural Networks, 119 (2019),162-177.*
- *Paper 2: Sengupta, N., McNabb, C. B., Kasabov, N., & Russell, B. R. (2018). Integrating Space, Time, and Orientation in Spiking Neural Networks: A Case Study on Multimodal Brain Data Modelling. IEEE Transactions on Neural Networks and Learning Systems, 29(11). doi:10.1109/TNNLS.2018.2796023*
- *Paper 3: Li, Jiawei; Liu, Jinyuan; Zhou, Shihua; Zhang, Qiang; Kasabov, Nikola, , "GeSeNet: A General Semantic-guided Network with Couple Mask Ensemble for Medical Image Fusion" , IEEE Transactions on Neural Networks and Learning Systems, DOI: https://doi.org/10.1109/TNNLS.2023.3293274, 21 July 2023.*

4. MDSE for predictive modelling of multisensory streaming data

- *Main: N.Kasabov, Chapter 19, Time-space, spiking neural networks and brain-inspired AI, Springer-Nature, 2019.*
- *Additional papers:*
- *H Liu, G Lu,Y Wang, N Kasabov, Evolving spiking neural network model for PM2.5 hourly concentration prediction based on seasonal differences: A case study on data from Beijing and Shanghai, Aerosol and Air Quality Research, vol.21, Issue 2, Feb. 2021, 200247, https://doi.org/10.4209/aaqr.2020.05.0247*
- *Laña I, Lobo JL, Capecci E, Del Ser J, Kasabov N, Adaptive long-term traffic state estimation with evolving spiking neural networks, Transportation Research Part C: Emerging Technologies **101**:126-144 2019, https://doi.org/10.1016/j.trc.2019.02.011*

5. MDSE for integrated audio-visual information processing

- *Main reading: N.Kasabov, Chapter 13, Time-space, spiking neural networks and brain-inspired AI, Springer-Nature, 2019.*
- *Additional papers:*

Paper 1: N. Kasabov et al, AVIS: a connectionist-based framework for integrated auditory and visual information processing. Inf. Sci. 133, 137–148 (2000)

Paper 2: Wysoski, S., L.Benuskova, N.Kasabov, Evolving Spiking Neural Networks for Audio-Visual Information Processing, Neural Networks, 23, 7, 819-835, 2013

Paper 3: N Kasabov, B Bhattacharya, D Patel, N Aggarwal, T Bankar, I AbouHassan, Cognitive Audio-Visual Associative Memories using Brain-inspired Spiking Neural Networks with Case Studies on Moving Object Recognition (IEEE Trans. Cognitive and Devel. Systems, 2023).

6. MDSE for integrating times series and text data in finance and economics (Ms Iman AbouHassan and Prof. Kasabov)

- *Paper: I AbouHassan, N Kasabov, V Jagtap, P Kulkarni, Spiking neural networks for predictive and explainable modelling of multimodal streaming data on the Case Study of Financial Time Series Data and on-line news, SREP, Springer-Nature, Sci Rep 13, 18367 (2023). https://doi.org/10.1038/s41598-023-42605-0*

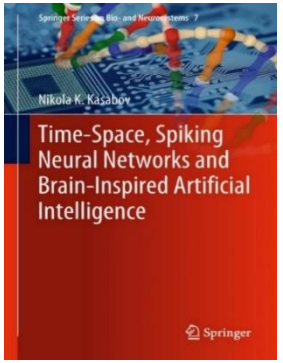
7. MDSE for integration of brain data and face image data for emotion recognition

- *Paper: C Tan; G Ceballos; N Kasabov; N Subramaniyam, FusionSense: Emotion Classification using Feature Fusion of Multimodal Data and Deep learning in a Brain-inspired Spiking Neural Network, Sensors (ISSN 1424-8220), MDPI Publisher, September 2020*

8. Revision of the course

nkasabov@aut.ac.nz

Lecture 5. MDSE for integrated audio-visual information processing



Main reading: N.Kasabov, Chapter 13, Time-space, spiking neural networks and brain-inspired AI, Springer-Nature, 2019.

Relevant papers:

N. Kasabov et al, AVIS: a connectionist-based framework for integrated auditory and visual information processing. Inf. Sci. 133, 137–148 (2000)

N Kasabov, B Bhattacharya, D Patel, N Aggarwal, T Bankar, I AbouHassan, Cognitive Audio-Visual Associative Memories using Brain-inspired Spiking Neural Networks with Case Studies on Moving Object Recognition (IEEE Trans. Cognitive and Devel. Systems, 2023).

Problems with multimodal audio-visual data modelling

- They are processed always together in the human brain, but rarely in AI systems
- They are of different time and space scales
  - granularity of integration
  - time alignment in the integration process (supporting revision)
  - real-time feedback
- They relate to other multimodal data and context, e.g. gestures, emotion, etc.
- Examples:
  - Speech recognition with lip reading
  - Person authentication using both speech and image
  - Moving object recognition
  - Chat bots for human to machine communication
  - Robot communication

Multi-modal person identification

Paper 1: N. Kasabov et al, AVIS: a connectionist-based framework for integrated auditory and visual information processing. Inf. Sci. 133, 137–148 (2000)

- The task is to identify persons from video information
- Integrating voice and face recognition
- Segmenting the video information
- A NN is trained on **synchronised** 105 image features and 78 speech features
- The recognition, when integrated input is used, is better than the one when either only speech or only image is used.

nkasabov@aut.ac.nz htpps://knowledgeengineering.ai

Multiple modality systems are better than single modality systems

- The test (out of sample) results (the % of correct classification rate) of the classification of the 29 test frames by PIAVI's subsystems (SS) and by persons (P).

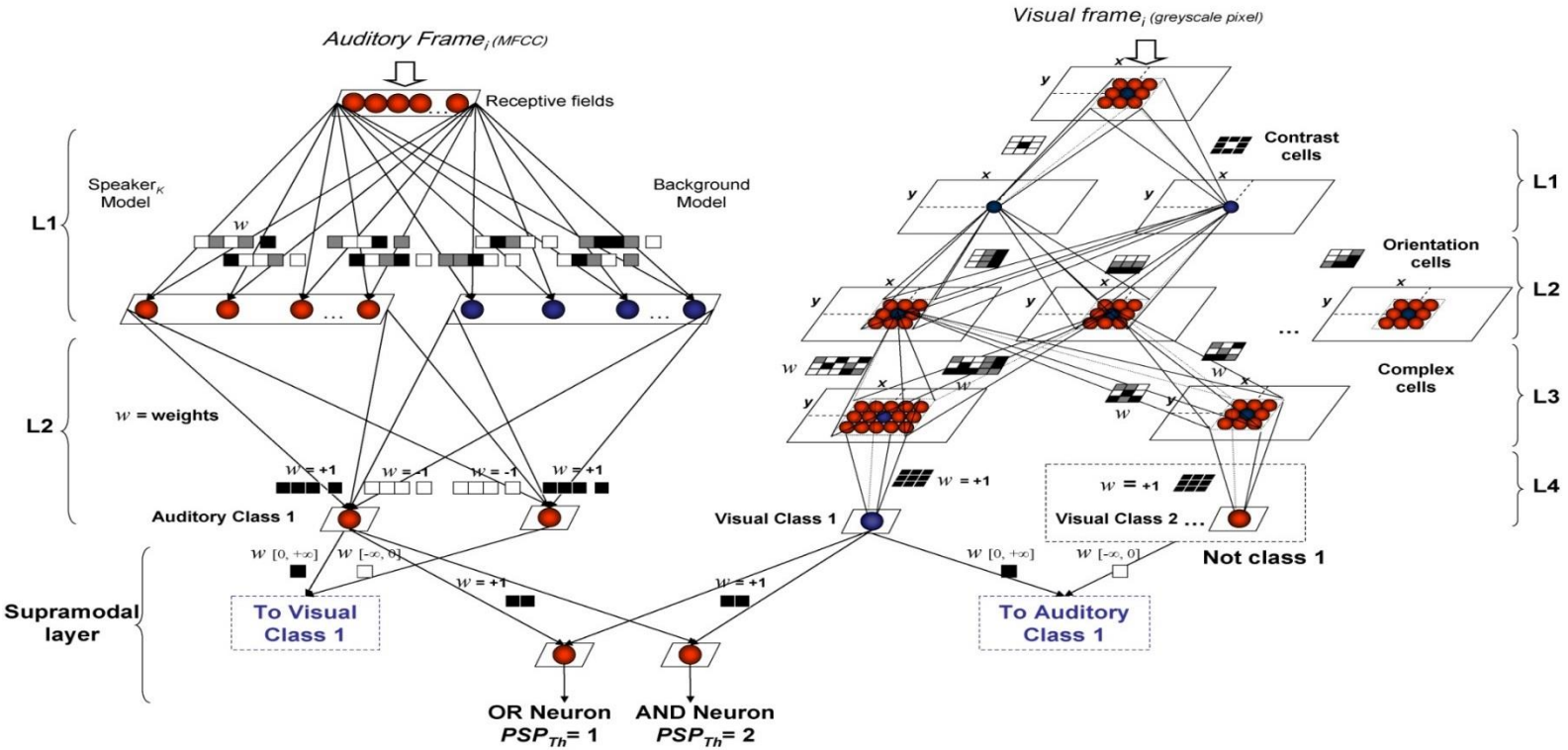
	P1	P2	P3	P4	%
Auditory SS only	7	5	3	9	60
Visual SS only	4	6	5	8	57.5
Early integration	7	8	3	8	67.5
Late integration	7	8	4	9	70

Audio-visual information processing for person authentication

-Convolutional layers

- Accumulation of spikes over time in the membrane potential

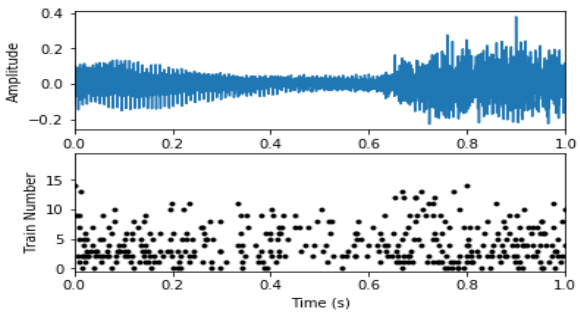
- Person authentication based on speech and face data

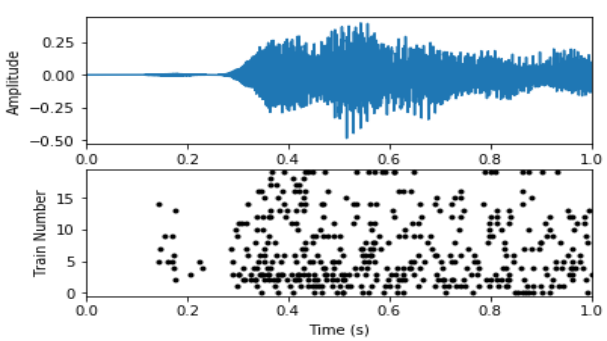


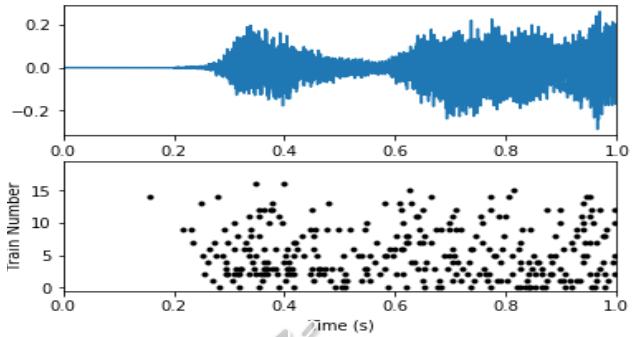
(Paper 2: Wysoski, S., L.Benuskova, N.Kasabov, Evolving Spiking Neural Networks for Audio-Visual Information Processing, Neural Networks, 23, 7, 819-835, 2013).

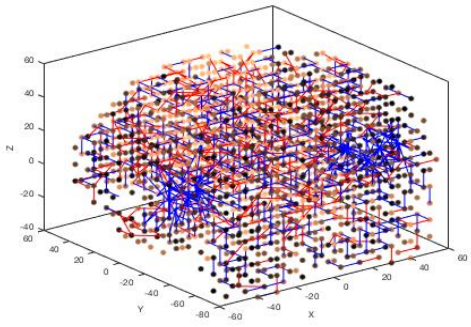
Deep learning and modelling of audio and visual and multimodal audio-visual data in BI-SNN (Chapter 13)

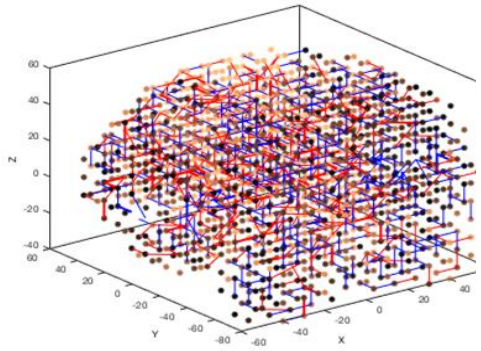
Using tonotopic, *stereo* mapping of sound and deep learning in NeuCube

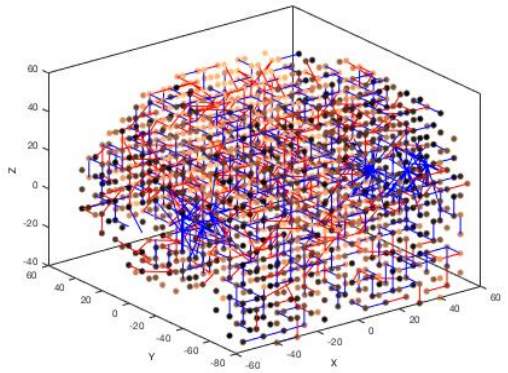










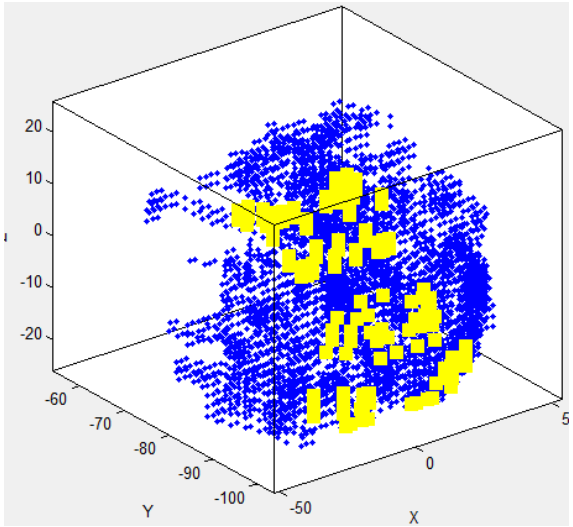


	Mozart	Bach	Vivaldi
Predicted 1	171	3	1
Predicted 2	9	176	1
Predicted 3	0	1	178

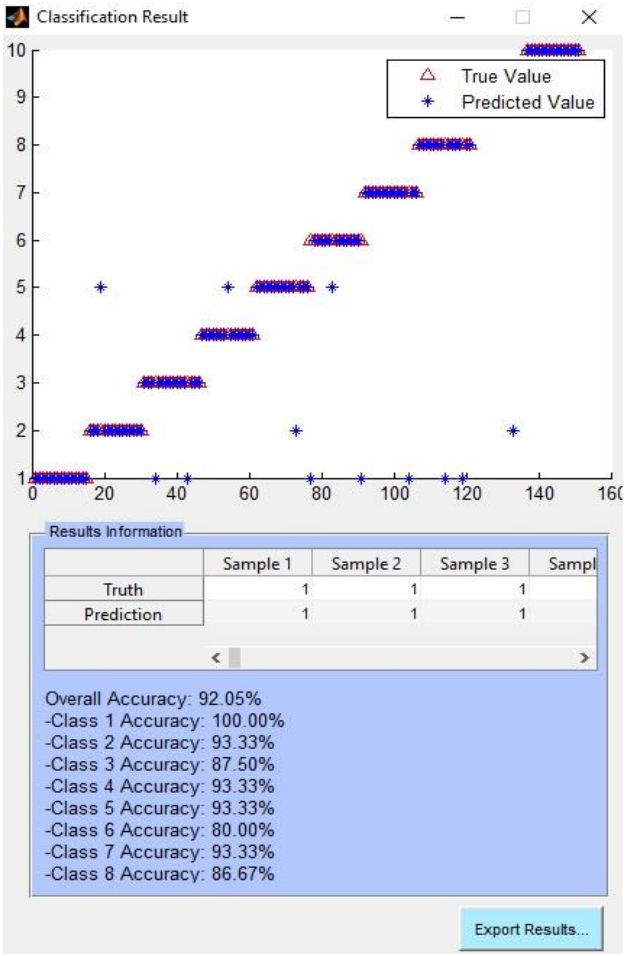
https://kedri.aut.ac.nz/R-and-D-Systems/audio-visual-data-processing-and-concept-formation#auditory

Deep learning and knowledge representation of moving objects using DVS and retinotopic mapping in NeuCube

30000 moving digits in 8 fonts and sizes from DVS MNIST



NeuCube with 4262 neurons from V1 and V2

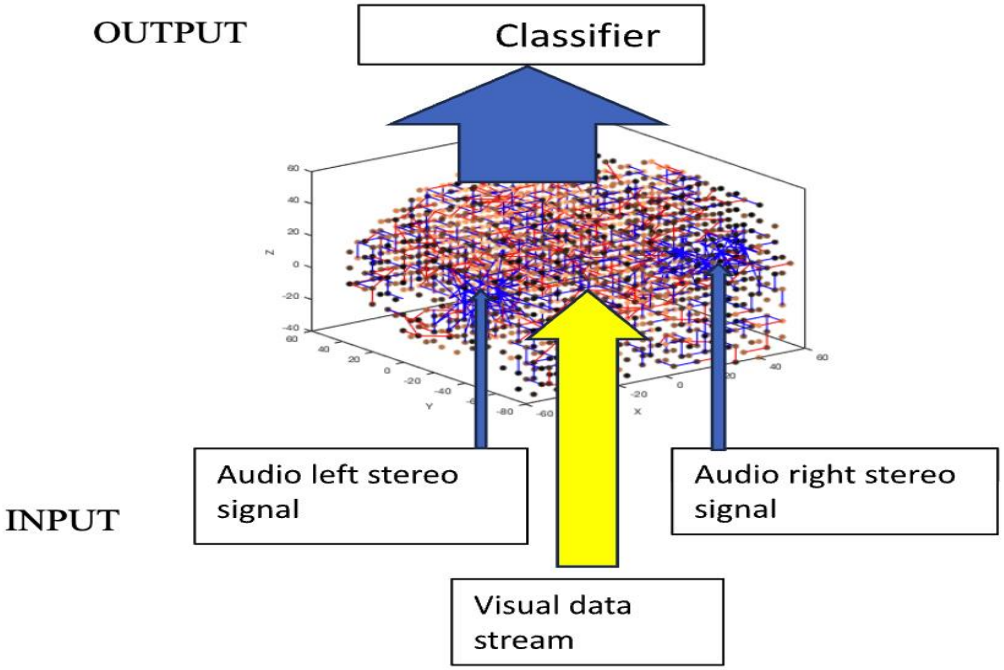


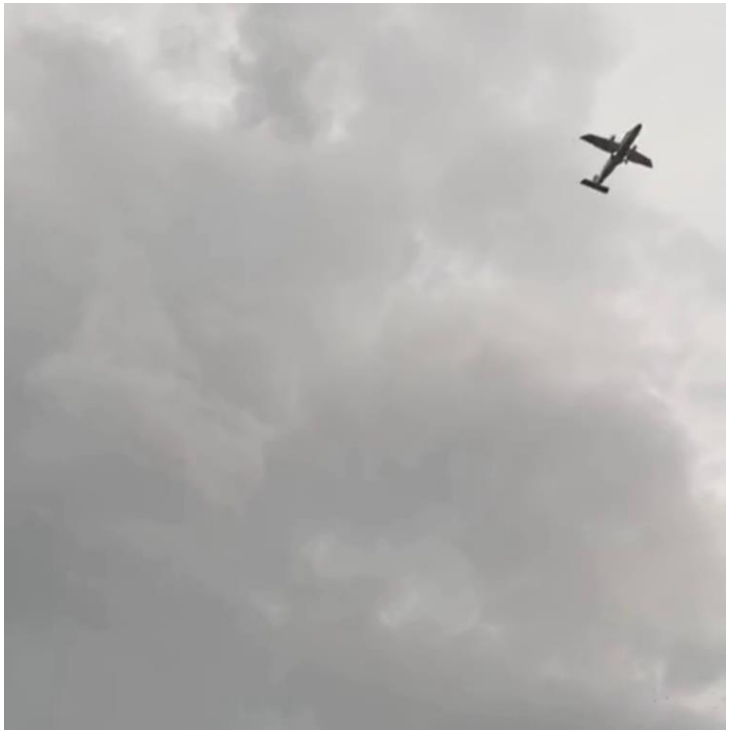
L.Paulin, A.Abbott, N.Kasabov, A retinotopic spiking neural network system for accurate recognition of moving objects using NeuCube and dynamic vision sensors, Frontiers of Comp. Neuroscience, 2018, doi:10.3389/fncom.2018.00042.

Multimodal audio-visual STAM:
Moving object recognition when missing modality

Paper 3: Kasabov, N., B.Bhattacharya, et al, AViAM-SNN: A Framework for Audio-Visual Associative Memories using Brain-inspired Spiking Neural Networks, Applied Soft Computing, 2023 (submitted)

- Training on both modalities
- Recall on only one modality
- Case studies on airplane and train recognition





nkasabov@aut.ac.nz