

# Midterm Paper Study – Coding in AI 2023

## Rules

1. Group Size: Each group consists of 2-3 people.
2. Bidding Points: Every group starts with 100 bidding points.
3. Topic Bids: You can bid on 1-3 topics.
4. Example Bids:
  - Bid for Paper 1 - 50 points.
  - Bid for Paper 2 - 49 points.
  - Bid for Paper 3 - 1 point.
5. Paper Assignment:
  - Each paper will be assigned to the group with the highest bid points.
  - If you do not win any bids, you will be assigned a paper that was not chosen by others.
  - In the event of two or more groups placing equal bids on one paper, the group that placed the bid first will be awarded that paper.
6. Bid Timing: The bid placement form will be open from 6:00 PM today and will close at midnight.
7. Result Announcement: Bidding results will be announced tomorrow.

## Midterm Paper

You have to write a review report based on your assigned research paper and present it in the class.

Your review report should consist of the following sections:

1. Abstract
2. Introduction
3. Theory
4. Experimental Design and Results
5. Conclusion
6. References

## Images

I1	<b>"ImageNet Classification with Deep Convolutional Neural Networks"</b> <ul style="list-style-type: none"><li>• Authors: Alex Krizhevsky et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
I2	<b>"Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks"</b> <ul style="list-style-type: none"><li>• Authors: Shaoqing Ren et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
I3	<b>"Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs"</b> <ul style="list-style-type: none"><li>• Authors: Liang-Chieh Chen et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
I4	<b>"Mask R-CNN"</b> <ul style="list-style-type: none"><li>• Authors: Kaiming He et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
I5	<b>"Show and Tell: A Neural Image Caption Generator"</b> <ul style="list-style-type: none"><li>• Authors: Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan</li><li>• Link: <a href="#">arXiv</a></li></ul>
I6	<b>"Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks"</b> <ul style="list-style-type: none"><li>• Authors: Alec Radford, Luke Metz, and Soumith Chintala</li><li>• Link: <a href="#">arXiv</a></li></ul>
I7	<b>"EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks"</b> <ul style="list-style-type: none"><li>• Authors: Mingxing Tan and Quoc V. Le</li><li>• Link: <a href="#">arXiv</a></li></ul>

## Video/Sequence Images

V1	<b>"You Only Look Once: Unified, Real-Time Object Detection"</b> <ul style="list-style-type: none"><li>• Authors: Joseph Redmon et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
V2	<b>"Fully-Convolutional Siamese Networks for Object Tracking"</b> <ul style="list-style-type: none"><li>• Authors: Luca Bertinetto et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
V3	<b>"SlowFast Networks for Video Recognition"</b> <ul style="list-style-type: none"><li>• Authors: Christoph Feichtenhofer et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
V4	<b>"Learning Spatiotemporal Features with 3D Convolutional Networks"</b> <ul style="list-style-type: none"><li>• Authors: Du Tran et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
V5	<b>"Temporal Convolutional Networks for Action Segmentation and Detection"</b> <ul style="list-style-type: none"><li>• Authors: Victor Escorcia et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
V6	<b>"Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset"</b> <ul style="list-style-type: none"><li>• Authors: Joao Carreira and Andrew Zisserman</li><li>• Link: <a href="#">arXiv</a></li></ul>
V7	<b>"Long-term recurrent convolutional networks for visual recognition and description"</b> <ul style="list-style-type: none"><li>• Authors: Jeff Donahue et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>

## Text

T1	<b>"Attention Is All You Need" (10)</b> <ul style="list-style-type: none"> <li>Authors: Ashish Vaswani et al.</li> <li>Link: <a href="#">arXiv</a></li> </ul>
T2	<b>"BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding" (9)</b> <ul style="list-style-type: none"> <li>Authors: Jacob Devlin et al.</li> <li>Link: <a href="#">arXiv</a></li> </ul>
T3	<b>"XLNet: Generalized Autoregressive Pretraining for Language Understanding"</b> <ul style="list-style-type: none"> <li>Authors: Zhilin Yang et al.</li> <li>Link: <a href="#">arXiv</a></li> </ul>
T4	<b>"ULMFiT: Universal Language Model Fine-tuning for Text Classification" (9)</b> <ul style="list-style-type: none"> <li>Authors: Jeremy Howard and Sebastian Ruder</li> <li>Link: <a href="#">arXiv</a></li> </ul>
T5	<b>"Sequence to Sequence Learning with Neural Networks" (8)</b> <ul style="list-style-type: none"> <li>Authors: Ilya Sutskever et al.</li> <li>Link: <a href="#">arXiv</a></li> </ul>
T6	<b>"Wabiq: A wikipedia-based thai question-answering system" (8)</b> <ul style="list-style-type: none"> <li>Authors: Noraset Thanapon et al.</li> <li>Link: <a href="#">sciencedirect.com</a></li> </ul>
T7	<b>"Attacut: A fast and accurate neural thai word segmenter" (8)</b> <ul style="list-style-type: none"> <li>Authors: Chormai Pattarawat et al.</li> <li>Link: <a href="#">arXiv</a></li> </ul>

## Speech

S1	<b>"Deep Speech: Scaling up end-to-end speech recognition"</b> <ul style="list-style-type: none"><li>• Authors: Awni Hannun et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
S2	<b>"Listen, Attend and Spell"</b> <ul style="list-style-type: none"><li>• Authors: Chan et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
S3	<b>"WaveNet: A Generative Model for Raw Audio"</b> <ul style="list-style-type: none"><li>• Authors: van den Oord et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
S4	<b>"Deep Voice: Real-time Neural Text-to-Speech"</b> <ul style="list-style-type: none"><li>• Authors: Serkan O. Arik et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
S5	<b>"Convolutional-recurrent neural networks for speech enhancement"</b> <ul style="list-style-type: none"><li>• Authors: Zhao et al.</li><li>• Link: <a href="#">IEEE</a></li></ul>
S6	<b>"Looking to Listen at the Cocktail Party: A Speaker-Independent Audio-Visual Model for Speech Separation"</b> <ul style="list-style-type: none"><li>• Authors: Luo et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>
S7	<b>"Augmenting Generative Adversarial Networks for Speech Emotion Recognition"</b> <ul style="list-style-type: none"><li>• Authors: Siddique Latif et al.</li><li>• Link: <a href="#">arXiv</a></li></ul>