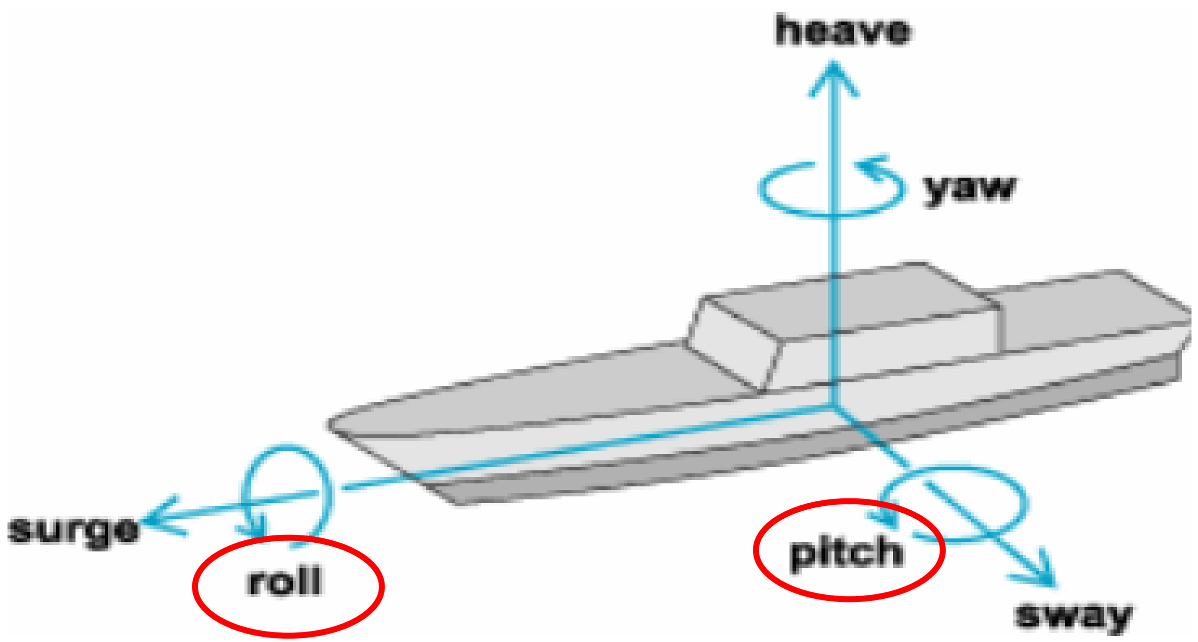
Meeting 19-11-2021

Deep learning for predicting ship motion from images

- Predict pitch and roll 30 seconds ahead
- CNN for image processing
- LSTM network for time series processing
- Autoencoder: pre-trained component in the models

Parameters



- Rotational
  - Yaw (human)
  - Pitch & Roll (wave)
- Translational
  - *Heave*, sway & surge
- Video/images
  - Camera:
    - Height, x-y-z rotations
    - Find optimum fps

Data preprocessing

- Pixel data normalized to [-1,1]
- Pitch and Roll normalized with absolute values [-90°,90°] to [-1,1]

Models

- CNN, LSTM networks, Autoencoders (performance)
- Images for input stacked or not
- **FC** - Fully connected layers
- **PR** - Pitch and Roll used as input
- **Encoder** – LSTM architecture used to encode input vector
- **img-encoder** – individual encoder to encode sequence of images
- **PR-encoder** – individual encoder to encode sequence of pitch & roll
- **Decoder** – LSTM architecture used to decode latent vector
- **MSE** – pitch and roll at 10th predicted second

CNN models

CNN stack FC model	CNN stack PR FC model	CNN PR FC model
• Predict sequence of seconds: higher MSE but justified • 10^{th} second MSE of p:28° and r:28°	• Pitch and roll now considered, increased accuracy • 10^{th} second MSE of p:21° and r:22°	• Seperate images as input: accuracy increase • Poor convergence • 10^{th} second MSE of p:11° and r:11°

LSTM CNN hybrid models

LSTM encoder decoder PR model

- Sequence of past to sequence of future frames
- 10th second MSE of p:30° and r:29°

CNN LSTM img-encoder PR-encoder decoder

- CNN extracts features, combined in vector for img-encoder
- Pitch and Roll separately encoded
- One latent vector is decoded
- 10th second MSE of p:3.5° and r:2.8°

LSTM CNN hybrid models

CNN LSTM encoder decoder images PR

- Combine img-encoder and pr-encoder for less complexity
- 10th second MSE of **p:2.9°** and **r:2.7°**

CNN LSTM encoder decoder images

- Pitch and roll dropped
- 10th second MSE of p:3.4° and r:2.3°

CNN LSTM images PR

- Combine all LSTM elements into one
- Reduce parameters, higher efficiency
- 10th second MSE of p:4.2° and r:4.1°

Optimizations

- ADAM: adaptive learning rate optimization
- MSE loss fuction
- Early stopping rules to prevent overfitting
- Hyperband algorithm for hyperparameter tuning

Future improvements

- Test new sequence to sequence models
  - GRU-CNN model: faster then LSTM (Gated Recurrent Unit)
- Test other hyperparameters (try keeping them as fixed as possible)
- Using real images/video: how do the results change
- Using more ship parameters/weather data (mainly wind & current) (available is blender)
- Clever data augmentation: extend trainingset (mirror images, add fog, rain)

Next presentaion

- The data split
  - Train/Test/Evaluation split
  - Test set is used for overfitting
  - Evaluation set is used for MSE
- Criteria to benchmark the models and compare them
  - MSE relative to maximum values for pitch and roll
    - 2.9° MSE is very high when the max value is only 10° roll
    - What are the maxima
  - Plot the results
- Architectures for the models

Sources

- Deep Learning For Predicting Ship Motion From Images – Nazar-Mykola Kaminskyi - https://github.com/Nazotron1923/Deep_learning_models_for_ship_motion_prediction_from_images/
- Real-Time Ship Motion Forecasting Using Deep Learning – 2021 – Mohammad Hasanur Rashid, Jing Zhang, Minghao Zhao - https://dl.acm.org/doi/abs/10.1145/3448734.3450923
- A Neural Conversational Model – 2015 - Oriol Vinyals, Quoc Le - https://arxiv.org/abs/1506.05869
- A Co-operative Hybrid Model For Ship Motion Prediction - R. Skulstad, G. Li, T. I. Fossen, T. Wang, H. Zhang - https://www.researchgate.net/publication/351397594_A_Co-operative_Hybrid_Model_For_Ship_Motion_Prediction