HOW TO ADDRESS UNCERTAINTY IN SMALLER, FASTER, MORE AGILE, YET SAFER DRONES?

ERDAL KAYACAN
AARHUS UNIVERSITY,
DEPARTMENT OF ENGINEERING
DIRECTOR OF ARTIFICIAL INTELLIGENCE IN ROBOTICS (AIR) LAB

IJCCI-ROBOVIS 2020 KEYNOTE
NOVEMBER 4, 2020
### WHO AM I?

#### Work experience

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
</table>
| Apr 2018-Current | Associate Professor, Aarhus University, *Department of Engineering*  
                     *Director of Artificial Intelligence in Robotics (AiR) Lab*      |
| Mar 2014-Mar 2018| Assistant Professor, Nanyang Technological University, Singapore  
                     *Department of Mechanical and Aerospace Engineering*             |
| Sep 2011-Mar 2014| Post doctoral researcher, KU Leuven, Belgium  
                     *Division of Mechatronics, Biostatistics and Sensors (MeBioS)* |

#### Education

<table>
<thead>
<tr>
<th>Date</th>
<th>Degree</th>
<th>Institution</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sep 2011</td>
<td>Ph.D. in Electrical and Electronics Engineering, Bogazici University</td>
<td>Istanbul</td>
<td></td>
</tr>
<tr>
<td>Jan 2016</td>
<td>M.Sc. in Systems and Control Engineering, Bogazici University</td>
<td>Istanbul</td>
<td></td>
</tr>
<tr>
<td>Jun 2003</td>
<td>B.Sc. in Electrical Engineering, Istanbul Technical University</td>
<td>Istanbul</td>
<td></td>
</tr>
</tbody>
</table>
## RESEARCH PROJECTS

<table>
<thead>
<tr>
<th>Date</th>
<th>Ongoing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec 2020 – Dec 2024</td>
<td>Reliable AI for Marine Robotics (ReMaRo)</td>
</tr>
<tr>
<td></td>
<td>by Horizon 2020 - H2020-MSCA-ITN-2020, European Union</td>
</tr>
<tr>
<td>Jan 2020 – Jan 2023</td>
<td>Open Deep Learning toolkit for Robotics (OpenDR)</td>
</tr>
<tr>
<td></td>
<td>by Robotics Core Technology ICT-10-2019-2020, European Union</td>
</tr>
<tr>
<td>Apr 2020 – Jun 2021</td>
<td>Smart Parking System for Vessels and Ports</td>
</tr>
<tr>
<td></td>
<td>by European Regional Development Fund</td>
</tr>
<tr>
<td>Jun 2020 – Jun 2021</td>
<td>Vision-based inspection navigation algorithm for ship inspection</td>
</tr>
<tr>
<td></td>
<td>by European Regional Development Fund</td>
</tr>
<tr>
<td>Jan 2021 – Jan 2022</td>
<td>Autonomous inspection of wind turbines using drones</td>
</tr>
<tr>
<td></td>
<td>by European Regional Development Fund</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>Directed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 2019-Dec 2019</td>
<td>Visualisation of Virtual Outcrops Using Aerial Robots</td>
</tr>
<tr>
<td></td>
<td>by Technical University of Denmark, Danish Hydrocarbon Research and Technology Centre</td>
</tr>
<tr>
<td>Mar 2018 - Mar 2019</td>
<td>Learning-based path planning of unmanned aerial vehicles with vision-based sensing</td>
</tr>
<tr>
<td></td>
<td>by Ministry of Education Academic Research Funding Tier 1</td>
</tr>
<tr>
<td>Jan 2016 - Apr 2018</td>
<td>Fuzzy neural network-based learning control of unmanned aerial vehicles</td>
</tr>
<tr>
<td></td>
<td>by ST Eng-NTU Corporation Laboratory</td>
</tr>
<tr>
<td>Jan 2014 - Apr 2018</td>
<td>Design of lightweight UAV for 3D Printing</td>
</tr>
<tr>
<td></td>
<td>by NRF Medium-Sized Centre</td>
</tr>
<tr>
<td>Jul 2015 - Dec 2017</td>
<td>Precise landing for unmanned aerial vehicles</td>
</tr>
<tr>
<td></td>
<td>by ST Eng-NTU Corporation Laboratory</td>
</tr>
<tr>
<td>Jul 2015 - Jan 2017</td>
<td>Quality Inspection and Assessment Robot (Quicabot)</td>
</tr>
<tr>
<td></td>
<td>by JTC Corporation - NRF Singapore</td>
</tr>
<tr>
<td>May 2014 - Mar 2017</td>
<td>Learning control algorithms for unmanned aerial vehicles</td>
</tr>
<tr>
<td></td>
<td>by Nanyang Technological University (Start up grant)</td>
</tr>
<tr>
<td>Mar 2015 - Aug 2017</td>
<td>Model predictive control-moving horizon estimation framework as applied to tilt rotor UAVs and its experimental evaluation</td>
</tr>
<tr>
<td></td>
<td>by Ministry of Education Academic Research Funding Tier 1</td>
</tr>
</tbody>
</table>
MY MOTIVATION FOR ROBOTICS
ESSENTIAL UNITS IN ROBOTS

Paradigm shift

Control
Planning
Perception

Understand
Perception
Adapt
Control
Planning
Interact
Learn

Paradigm shift

ERDAL KAYACAN
NOVEMBER 4 2020
IJCCI-ROBOVIS 2020 KEYNOTE

HOW TO ADDRESS UNCERTAINTY IN SMALLER, FASTER, MORE AGILE, YET SAFER DRONES?
FIXED WING AND MULTI-ROTOR UAVS

Multi-rotor UAVs

Fixed-wing vertical take-off (VTOL)

Airbus electrical aerial taxi

Uber aerial taxi designs

Volocopter
**AUTONOMY VERSUS SIZE**

Source: Prof. Vijay Kumar, ICRA 2019 Keynote Speech
BATTERY TECHNOLOGY

Source: Battery technology

Fat (Wh/kg) = 10,000

Source: Prof. Vijay Kumar, ICRA 2019 Keynote Speech
WHAT KIND OF ONBOARD SENSORS?

Why is vision more popular?

- **GPS**: unreliable in urban environment, lost in indoor
- **IMU**: low accuracy, accumulated error
- **Laser**: limited range, heavy, expensive, high computing cost
- **Camera**
WHY LEARNING CONTROL?
WHAT HAPPENS IF YOUR CONTROLLER DOES NOT WORK?
"Given, for one instant, an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it an intelligence sufficiently vast to submit these data to analysis it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom. For it, nothing would be uncertain and the future, as the past would be present to its eyes”.

— Pierre-Simon Laplace
WHY MODEL-FREE LEARNING CONTROL?
LEARNING CONTROL USING ANNS

Artificial Neural Networks:

- The capability of handling uncertain information
- and the capability learning from input-output data

For tuning the parameters of ANNs:
1. The gradient based algorithm (include partial derivatives, the convergence speed, local minimum, etc)
2. Evolutionary approaches (the stability, computational burden)
3. Sliding mode control theory-based algorithms?

Theorem: If the adaptation laws for T2FNN parameters are chosen as:

\begin{align*}
e_{11} &= -\beta_1 \frac{\sigma_1}{e-e_{11}} \text{sgn}(u_c) \quad \text{and} \quad \dot{e}_{21} = -\beta_1 \frac{\sigma_2}{e-e_{21}} \text{sgn}(u_c) \\
\dot{\sigma}_{11} &= -\beta_1 \frac{\sigma_1}{(e-e_{11})^2} \text{sgn}(u_c) \quad \text{and} \quad \dot{\sigma}_{21} = -\beta_1 \frac{\sigma_2}{(e-e_{21})^2} \text{sgn}(u_c) \\
\dot{\tau}_{11} &= -\beta_1 \frac{\tau_1}{(e-e_{11})^2} \text{sgn}(u_c) \quad \text{and} \quad \dot{\tau}_{21} = -\beta_1 \frac{(\tau_2)^3}{(e-e_{21})^2} \text{sgn}(u_c) \\
\dot{i}_{1j} &= -\alpha \frac{q \dot{W}_{1j} + (1-q) \dot{W}_{2j}}{(q \dot{W} + (1-q) \dot{W})^T (q \dot{W} + (1-q) \dot{W})} \text{sgn}(u_c) \\
\dot{\alpha} &= \gamma_1 |u_c| - \nu \gamma_1 \alpha
\end{align*}

then, given an arbitrary initial condition $u_c(0)$, the learning error $u_c(t)$ will converge firmly to zero during a finite time $t_f$.


SOME APPLICATIONS
Novel Levenberg-Marquardt Based Learning Algorithm for Unmanned Aerial Vehicles

Andriy Sarabakha, Nursultan Imanberdiyev, Erdal Kayacan, Mojtaba Ahmadieh Khanesar, and Hani Hagras

March 2017
A Fast Learning Control Strategy for Unmanned Aerial Manipulators

Nursultan Imanberdiyev and Erdal Kayacan

School of Mechanical and Aerospace Engineering,
Nanyang Technological University, Singapore

Department of Engineering,
Aarhus University, Denmark

May 2018
**ONLINE TUNING OF DEEP NEURAL NETWORKS**

(a) Offline pre-training of DFNN by conventional controller.

(b) Online training of DFNN by FLS.


Andriy Sarabakha and Erdal Kayacan, Online Deep Fuzzy Learning for Control of Nonlinear Systems Using Expert Knowledge, Trans. on Fuzzy Systems (Accepted for publication)
Online Deep Fuzzy Learning for Control of Nonlinear Systems Using Expert Knowledge

Andriy Sarabakha and Erdal Kayacan

School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore
Department of Engineering, Aarhus University, Denmark

November 2018
AUTONOMOUS DRONE RACING
GATENET: EFFICIENT DEEP NEURAL NETWORK FOR GATE PERCEPTION IN AUTONOMOUS DRONE RACING
Problem: Can we fly better – faster and safer – than a skilled drone pilot using onboard processing?

"The AlphaPilot Innovation Challenge offers a chance for teams to master autonomous flight and win more than $2,000,000 in cash prizes!"
How can we fly faster and safer than a human pilot?

• Sense-Plan-Act as fast/accurate as possible
• perception and action loops are coupled
• Lack of modeling
• Reduce latencies: faster sensors and/or algorithms
• Additional problems because of ultra fast speed: motion blur
MOTIVATION

- Technologies used in drone racing can help to significantly increase the airborne time and coverage of a mission.
- Fast and reliable gate perception plays a vital role in the overall success of the race.
- **Problem definition**: Design a perception system to perceive unknown gates in a cluttered environment, that is robust to noisy backgrounds and gate pose variations.

A drone racing competition with human pilots (courtesy of Drone Racing League)
Related Work

- Traditional CV methods tend to fail in complex background with varying lighting conditions, occlusion, and blurriness ([de Croon, 2016], [Jung 2017], [Li, 2020])
- Deep neural network (DNN) methods ([Jung, 2018], [Kaufmann, 2019]) perform more accurate and robust, but they use large network size (0.5 - 7.8 M params) and have low inference rates.
RELATED WORK

- A hybrid method that segment the 4 corners and use a DNN to associate them to a gate [Foehn, 2020], but it is expensive, and not robust if gates are too close.
- End-to-end planning methods: map image input directly to control input / actions to plan the robot ([Muller, 2018], [Camci, 2019], [Zhou, 2019]). However the actions of the robot is hard to verify.
We propose a novel DNN, GateNet, that are:

- Accurate with equal or better performance in a wide range of different scenarios
- Has high inference rate (~ 60 Hz) on a real onboard processor
- Robust to gate pose changes and background disturbances
- Work with fisheye camera lenses with wide field-of-view (FOV) and can detect multiple gates
**METHODOLOGY**

- GateNet unifies predictions of (i) center, (ii) orientation, and (iii) distance of gates in a single neural network.
- The design and the number of parameters of the network enable high inference speed that is crucial for drone racing.
- It is applicable to other problem domains requiring object detection during an agile flight.
AU-DR DATASET

We propose a public dataset that focuses on gate perception in drone racing to help train and benchmark gate perception methods, available on like at: https://github.com/open-airlab/GateNet.git.

We explicitly label the images according to gate layouts, including images with (i) single gate, (ii) multiple gates, (iii) occlusion, (iv) partially observable gates, (v) too far gates, and (vi) too close gates.
TRAINING

- Target samples for training are prepared according to the spatial layout of gates in an input image by dividing the image into grids, and assign a confidence value for each square.
- The network is trained to minimize a multi-part objective function.
IN-HOUSE RACING DRONE

- In-house Racing Frame: 250 mm
- Autopilot: PX4 / Snapdragon Flight 820
- High frequency RGB Camera (up to 100 Hz)
- Intel Tracking camera T265
- Mass: < 1 kg, Thrust-to-weight ratio: ~ 4.0
EVALUATION AND RESULTS

- GateNet is superior in gate distance prediction compared to other baselines, and has better performance for gate orientation prediction in challenging cases such as partially observable gates or gates that are too close to the camera.
- Only ADRNet is better to find object centers than GateNet, but ADRNet uses much larger number of parameters, and cannot predict gate distance and orientation.

<table>
<thead>
<tr>
<th>method</th>
<th>#</th>
<th>fps</th>
<th>single gate</th>
<th>multiple gate</th>
<th>occlusion</th>
<th>partial</th>
<th>too close</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$E_c$ $E_d$ $E_\theta$</td>
<td>$E_c$ $E_d$ $E_\theta$</td>
<td>$E_c$ $E_d$ $E_\theta$</td>
<td>$E_c$ $E_d$ $E_\theta$</td>
<td></td>
</tr>
<tr>
<td>ADRNet [3]</td>
<td>2.5M</td>
<td>14</td>
<td>0.04 n/a n/a</td>
<td>0.07 n/a n/a</td>
<td>0.05 n/a n/a</td>
<td>0.04 n/a n/a</td>
<td></td>
</tr>
<tr>
<td>ADRNet-mod</td>
<td>2.5M</td>
<td>14</td>
<td>0.04 0.14 n/a</td>
<td>0.15 0.35 0.10</td>
<td>0.12 0.34 0.08</td>
<td>0.09 0.22 0.07</td>
<td></td>
</tr>
<tr>
<td>DroNet-mod</td>
<td>478K</td>
<td>22</td>
<td>0.16 0.18 0.10</td>
<td>0.50 0.60 0.42</td>
<td>0.37 0.45 0.29</td>
<td>0.13 0.17 0.09</td>
<td></td>
</tr>
<tr>
<td>Morales et al. [13]</td>
<td>1.1M</td>
<td>19</td>
<td>0.21 0.31 n/a</td>
<td>0.58 1.22 n/a</td>
<td>0.39 0.87 n/a</td>
<td>0.22 0.78 n/a</td>
<td></td>
</tr>
<tr>
<td>Darknet-mod</td>
<td>7.8M</td>
<td>9</td>
<td>0.10 0.12 0.05</td>
<td>0.20 0.25 0.12</td>
<td>0.14 0.24 0.09</td>
<td>0.15 0.15 0.08</td>
<td></td>
</tr>
<tr>
<td>GateNet</td>
<td>32K</td>
<td>57</td>
<td>0.05 0.10 0.05</td>
<td>0.16 0.24 0.12</td>
<td>0.15 0.39 0.12</td>
<td>0.07 0.14 0.06</td>
<td>0.05 0.11 0.07</td>
</tr>
</tbody>
</table>
REAL-WORLD EXPERIMENTS

- We demonstrate the effectiveness of the perception system on a fully-autonomous quadrotor system that flies on previously-unknown track with tight turns in a small environment.
- The prediction results of GateNet is used to back-project multiple gates’s poses on 3D world frame. Extended Kalman Filters are employed to keep the predicted poses stable on a Global map.
- The drone will plan and complete the track autonomously thanks to the global map.
MOTION PLANNING ARCHITECTURE

- Global planner: uses the perceived global gate map to generate a global trajectory through multiple gates.
- Local Replanner: Using a high-frequency receding horizon local planner to replan when (i) the next gate appears, (ii) there are changes in gate estimation.
REAL-WORLD EXPERIMENTS
EVALUATION OF ROBUSTNESS

- The drone completes a race track with multiple gates where the gates are close to each other (inter-distance ~ 4.5m) at 2 m/s
- GateNet can handle position and orientation changes in-between each lap with reasonable success rates up to 2m and 30°
- It is robust with background perturbations of human movements or gate piece disturbances.
ANOMALY DETECTION USING DRONES
Motivation

Taxonomy of anomalies:
- Point anomalies
Motivation

Taxonomy of anomalies:
- Point anomalies
- Collective anomalies
Motivation

Taxonomy of anomalies:
- Point anomalies
- Collective anomalies
- Contextual anomalies
Motivation

Taxonomy of anomalies:
- Point anomalies
- Collective anomalies
- Contextual anomalies
Motivation

Taxonomy of anomalies:
- Point anomalies
- Collective anomalies
- Contextual anomalies
Motivation

Taxonomy of anomalies:
- Point anomalies
- Collective anomalies
- Contextual anomalies
Motivation

Taxonomy of anomalies:
- Point anomalies
- Collective anomalies
- Contextual anomalies

Contextual anomalies are challenging:
- Requires a notion of context
- Normal region keeps change
Motivation

- Contextual anomaly detection is challenging for aerial surveillance.

- Anomalies can depend on time and location.

- UAVs can be deployed to a surveillance system.

- A UAV overcomes significant limitations of current video-based surveillance systems.
Contribution

CADNet: Deep neural network-based **context-aware anomaly detection method** for aerial traffic surveillance with a UAV.

In our experiments:
- Compare CADNet with state-of-the-approaches:
  - Autoencoder
  - One-class Support Vector Machine
  - Generative Adversarial Networks

- Analysis of the effect of contextual attributes (time and location) to anomaly detection performance
The method: CADNet

input frame $f$

object detector

FE Block
FC Block

context subnetwork

encoder

$\mu$
$h_1$
$\Sigma$
$h_2$
$z$
$h_3$
$\hat{x}$

decoder

Skip connection for $\hat{x}$

Skip connection for $c$

Skip connection for $x$

object detector output layer

$\sum z$
$h_3$
(32)
(128)
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 128$
Conv Layer $p \times 3 \times 64$
Conv Layer $p \times 3 \times 32$
Conv Layer $p \times 3 \times 16$
Conv Layer $p \times 3 \times 16$
# Experiments and Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CADNet</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>0.00</td>
<td>91.2%</td>
<td>86.6%</td>
</tr>
<tr>
<td>Autoencoder</td>
<td>n/a</td>
<td>n/a</td>
<td>yes</td>
<td>0.02</td>
<td>17.1%</td>
<td>21.4%</td>
</tr>
<tr>
<td>One-class SVM</td>
<td>n/a</td>
<td>n/a</td>
<td>yes</td>
<td>0.09</td>
<td>72.3%</td>
<td>69.2%</td>
</tr>
<tr>
<td>GAN</td>
<td>n/a</td>
<td>n/a</td>
<td>yes</td>
<td>0.03</td>
<td>48.3%</td>
<td>51.2%</td>
</tr>
</tbody>
</table>

### Ablation experiments follow below

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CADNet-wo-gps-time</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>0.00</td>
<td>90.2%</td>
<td>73.2%</td>
</tr>
<tr>
<td>CADNet-wo-skip</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>0.17</td>
<td>37.7%</td>
<td>35.0%</td>
</tr>
<tr>
<td>CADNet-wo-skip-c</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>0.00</td>
<td>88.5%</td>
<td>76.9%</td>
</tr>
<tr>
<td>CADNet-wo-skip-m</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>0.17</td>
<td>42.3%</td>
<td>32.9%</td>
</tr>
<tr>
<td>CADNet-wo-time</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>0.00</td>
<td>91.6%</td>
<td>54.4%</td>
</tr>
<tr>
<td>CADNet-wo-gps</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>0.00</td>
<td>88.7%</td>
<td>63.8%</td>
</tr>
<tr>
<td>CADNet-wo-frame</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>0.00</td>
<td>88.0%</td>
<td>78.6%</td>
</tr>
</tbody>
</table>

CADNet has better performance compared to the baselines.
Experiments and Results
DRONE MOVIE DIRECTORS
WHY IS FILMING DIFFICULT?

Drone filming

High-level decision making

Low-level decision making

Perception-vision

Perception-mapping
DEEP Q-LEARNING

Model-free solver using action-value function $Q$:

$$Q(c_t, a_t) = \sum_{t' = t}^T \mathbb{E}_{\pi, \theta} \left[ R_{act}(v(c_{t'}, a_{t'})) | c_t, a_t \right]$$

Deep Q function approximator:

Two approaches for reward function:

1) Hand-crafted reward:
   a) Actor’s presence ratio
   b) Shot angle
   c) Shot type duration
   d) Collision

2) Reward from human preferences:
EVALUATING THE LEARNED ARTISTIC BEHAVIOURS

Table 14: Average normalized score of video clips between 0 (worst) and 10 (best).

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>Scene 1</th>
<th>Scene 2</th>
<th>Scene 3</th>
<th>Scene 4</th>
<th>Scene 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand-crafted reward</td>
<td>8.2</td>
<td>10.0</td>
<td>5.3</td>
<td>9.3</td>
<td>7.7</td>
<td>8.7</td>
</tr>
<tr>
<td>Human reward</td>
<td>7.1</td>
<td>5.0</td>
<td>9.0</td>
<td>6.0</td>
<td>7.7</td>
<td>8.0</td>
</tr>
<tr>
<td>Back shot</td>
<td>3.8</td>
<td>4.0</td>
<td>4.7</td>
<td>4.3</td>
<td>4.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Random</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>0.3</td>
<td>0.7</td>
<td>1.3</td>
</tr>
</tbody>
</table>
## SUMMARY

A COMPLETE CINEMATOGRAPHY PIPELINE

- **Artistic reasoning**
- **Visual detection**
- **Mapping**
- **Motion planning**

### Our algorithm works with different targets

![Circling shot with bicycle x3](image)

- Shot type alterations make the video interesting.
- Keeping the actor in view is the most important criterion.
- Hand-crafted reward policy is the most exciting while the human reward policy is smooth.
WHAT IS NEXT?
CONCLUSIONS AND FUTURE PERSPECTIVES

- Not much room for further improvement for design (excluding fixed-wing VTOLs)
- Fast and nonlinear controllers are needed
- Perception and control cannot be (most of the time) separated
- Model-based and model-free methods can be used simultaneously
- Energy consumption is a big problem
HOW TO ADDRESS UNCERTAINTY IN SMALLER, FASTER, MORE AGILE, YET SAFER DRONES?

ERDAL KAYACAN
NOVEMBER 4 2020
IJCCI - ROBOVIS 2020 KEYNOTE

AIR-LAB TEAM

Erdal Kayacan
Ilker Bozcan, PhD candidate
Huy Pham, PhD candidate
Halil Ibrahim Ugurlu, PhD candidate
Jonas le Fevre Sejersen, Research Assistant
Rui Pimentel de Figueiredo, Post doc
WE MISS YOU
NOT EVERYTIME EVERYTHING GOES WELL...
erdal@eng.au.dk
http://www.erdal.info
https://www.youtube.com/erdalkayacan
https://www.linkedin.com/in/erdalk/