Neural Networks for Feature Extraction from Wendelstein 7-X Infra-Red Images

En Route towards Heat Load Control

Böckenhoff, Blatzheim, Niemann, Pisano, Hölbe, Jakuboswki and the W7-X Team
• Dynamic plasma properties ...

• Radial axis shift \[ \Delta R(t) = f(n, T) \]

• Rotational transform \[ \iota(t) = c \cdot \frac{l_{tor}}{\Theta} + \iota_{CF} \]
Motivation

• Dynamic plasma properties …
  • Radial axis shift \( \Delta R(t) = f(n, T) \)
  • Rotational transform \( \iota(t) = c \cdot \frac{i_{tor}}{\Theta} + \iota_{CF} \)

\[ \iota_{CF} = f(I_A, I_B, \ldots) \]
Motivation

• Dynamic plasma properties ...
  • Radial axis shift $\Delta R(t) = f(n, T)$
  • Rotational transform $\iota(t) = c \cdot \frac{l_{tor}}{\Theta} + \iota_{CF}$

\[ I_{tor} = l_{bs} \cdot e^{-t/\tau} \]
Motivation

• Dynamic plasma properties ...
  • Radial axis shift \[ \Delta R(t) = f(n, T) \]
  • Rotational transform \[ \iota(t) = c \cdot \frac{I_{\text{tor}}}{\Theta} + \iota_{CF} \]

• ... affect the PFC heat load
  ➢ Potential excessive heat loads \( q > q_d \)
  ➢ Altered performance (detachment, pumping, ...)

Avoid

Optimize / Control
Motivation

• Dynamic plasma properties ...
  • Radial axis shift \( \Delta R(t) = f(n, T) \)
  • Rotational transform \( \iota(t) = c \cdot \frac{I_{\text{tor}}}{\Theta} + \iota_{CF} \)

• ... affect the PFC heat load
  ➢ Potential excessive heat loads \( q > q_d \)
  ➢ Altered performance (detachment, pumping, ...)

➢ Long term objective: Real time heat load control
  ➢ PFC integrity
  ➢ Performance optimization
Motivation

• Dynamic plasma properties ...
  • Radial axis shift \( \Delta R(t) = f(n, T) \)
  • Rotational transform \( \iota(t) = c \cdot \frac{I_{\text{tor}}}{\Theta} + \iota_{CF} \)

• ... affect the PFC heat load
  ➢ Potential excessive heat loads \( q > q_d \)
  ➢ Altered performance (detachment, pumping, ...)

➢ Long term objective: Real time heat load control
  ➢ PFC integrity
  ➢ Performance optimization

• Missing real time estimation of \((\iota, \Delta R)\)
  ➢ Real time feature extraction \((\iota, \Delta R)\) diagnostic as add-on to IR-Cameras
Introduction

1. \( \iota \) reconstruction from limiter heat loads
2. \( \iota, \Delta R \) reconstruction from divertor heat loads
3. Conclusion / Outlook

Focus:
- Training Data generation and investigation
- NN input optimization
- Architecture experiments
OP 1.1

$l$ scan with first limiter plasmas

$l = ...$
OP 1.1 Setup
Data Set(s)

- \( \zeta \) Scan („Config. J“ – Op1.1)
Data Set(s)

• $\iota$ Scan ("Config. J" – Op1.1)
  • Sparse Experimental Data:
    - Only 6 distinct magnetic configurations

Infrared

Data Set Size

|\mathbb{I}| = 319

|\mathcal{S}| = 3993
Data Set(s)

• $t$ Scan ("Config. J" – Op1.1)
  • Sparse Experimental Data:
    ➢ Only 6 distinct magnetic configurations

Infrared

\[ \mathbb{I} \]

4

... \[ I_B \]

\[ t \]

Synthetic

\[ S \]

1000

FDL Parameters

\[ n_{tot} = 25 \times 10^3 \]
\[ \lambda = 0.1 \text{ m} \]
\[ v = 1.4 \times 10^5 \text{ m s}^{-1} \]
\[ D_\perp = 1 \text{ m}^2 \text{ s}^{-1} \]

Data Set Size

\[ |\mathbb{I}| = 319 \]
\[ |S| = 3993 \]
Data Set(s)

• \( \iota \) Scan ("Config. J" – Op1.1)
  • Sparse Experimental Data:
    ➢ Only 6 distinct magnetic configurations
  • Same trend but systematic differences in Simulations vs. Experiment

<table>
<thead>
<tr>
<th>Data Set Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ir</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

\[ |\iota| = 319 \]
\[ |S| = 3993 \]

\( n_{tot} = 25 \times 10^3 \)
\( \lambda = 0.1 \text{ m} \)
\( v = 1.4 \times 10^5 \text{ m s}^{-1} \)
\( D_\perp = 1 \text{ m}^2 \text{ s}^{-1} \)
Reconstruction of $\tau$ ($I_B$ as proxy) with NN from IR

First Approach:

• NN Training set: $\mathbb{I}$
Reconstruction of $\tau$ ($I_B$ as proxy) with NN from IR

First Approach:
- NN Training set: \[ \mathcal{X} \]

Lessons:
- Poor performance
- (too) few data

[Bockenhoff, Blatzheim et. al. NF 2018]
Performance Optimization: Synthetic Data

Optimization:

• NN Training set: $\mathcal{S}$
Performance Optimization: Synthetic Data

Optimization:

• NN Training set: $\mathcal{S}$
Performance Optimization: Synthetic Data

Optimization:
• NN Training set: $\mathbb{S}$

Lessons:
• Worse performance
• Different features in Sim. & Exp.
• NN „overfits“ towards Sim.

[Böckenhoff, Blatzheim et. al. NF 2018]
Performance Optimization: Mixing

Optimization:
- NN Training Set: $\mathcal{M} = \mathcal{S} \cup \mathcal{I}$
- Architecture & Preprocessing:

➢ Convolutional Neural Network (CNN)

Relative error < 5.4%
Optimization:

- NN Training Set: $\mathcal{M} = \mathcal{S} \cup \mathcal{I}$
- Architecture & Preprocessing:
  - Convolutional Neural Network (CNN)

Relative error < 5.4%

[Blatzheim, Böckenhoff et al. NF 2019]
Performance Optimization: Generative Adversarial Network

Optimization:

• **NN Training Set:** \( M = S \cup I \)

• **Architecture & Preprocessing:**
  - Convolutional Neural Network (CNN)
  - Generative adversarial neural network

\[ \text{Relative error} < 2.3\% \]

[Blatzheim, Böckenhoff et. al. submitted 2019]
OP 1.2

\( i, \Delta R \) scan with simulated divertor heat loads
OP 1.1 Setup

IRCAM

module 2

module 3

IRCAM

module 1

module 4

module 5
Divertor mapping

- Following physics and engineering constraints
- CNN Input
- Optimized for MC based simulations
  - Factor 20 more training data per time with same statistical significance

[Ref 3]
[Böckenhoff, Blatzheim et. al. NF 2019]
Simulated Data Set

- ~ 30000 Field Line Diffusion Simulations
Simulated Data Set

• ~ 30000 Field Line Diffusion Simulations
Simulated Data Set

\[ P_{\text{conv}} = 5 \text{MW} \]

(a) \((I_A, I_B) = (0.00, 0.00)\) i.e. standard reference

(b) \((I_A, I_B) = (0.25, 0.25)\) i.e. low iota reference

(c) \((I_A, I_B) = (-0.23, -0.23)\) i.e. high iota reference

(d) \((I_A, I_B) = (0.10, -0.20)\) i.e. inward shifted reference

(e) \((I_A, I_B) = (-0.14, 0.14)\) i.e. outward shifted reference

[Ref 3]
Overload Evaluation

• Evaluation of PFC integrity
• Reward function for Reinforcement Learning

[Böckenhoff, Blatzheim et al. NF 2019]
\[ \psi(t) = c \cdot I_{\text{tor}} \Theta + \psi_{\text{CF}} \] [Ref 3]

Toroidal Current Development

(a)

\[ P_{\text{conv}} = 8 \text{ MW} \]
\[ \iota(t) = c \cdot \frac{I_{\text{tor}}}{\Theta} + \iota_{CF} \]

Mimicked evolution of \( I_{\text{tor}} \) up to 44 kA

[Hölbe et al. NF 2016]
Toroidal Current Development

\[
\prod_i \Pr(X \leq n_{\text{crit},i} | n_i)
\]

\[I_A, I_B\]

\[P_{\text{conv}} = 8 \text{ MW}\]

[Ref 3]

[Böckenhoff, Blatzheim et. al. NF 2019]
Lessons

• Sweep can possibly mitigate overload arising from $I_{tor}$ evolution
### Experiments in NN Depth and Input Parametrization

<table>
<thead>
<tr>
<th>NN</th>
<th>Trainable Parameters</th>
<th>Learning Rate</th>
<th>Batch Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF − FC(_{ECl})</td>
<td>23983</td>
<td>0.0005</td>
<td>25</td>
</tr>
<tr>
<td>FF − FC(_{PBI})</td>
<td>270543</td>
<td>0.0010</td>
<td>25</td>
</tr>
<tr>
<td>CNN</td>
<td>63671</td>
<td>0.0005</td>
<td>25</td>
</tr>
<tr>
<td>DCNN</td>
<td>89073</td>
<td>0.0001</td>
<td>100</td>
</tr>
<tr>
<td>DINN</td>
<td>313071</td>
<td>0.0001</td>
<td>100</td>
</tr>
<tr>
<td>IRNN</td>
<td>804804</td>
<td>0.0001</td>
<td>100</td>
</tr>
</tbody>
</table>

- **rmse\(_{total}\)**
  - FF − FC\(_{ECl}\)
  - FF − FC\(_{PBI}\)
  - CNN
  - DCNN
  - DINN
  - IRNN

- **Training Time (\(10^3\) s)**

- cross validation set 1
- cross validation set 2
- cross validation set 3
- cross validation set 4
- cross validation set 5
- Training Time

[Blatzheim, Böckenhoff et. al. submitted to NF 2019]
(\( \tau, \Delta R \)) - Reconstruction

Experiments in NN Depth and Input Parametrization

<table>
<thead>
<tr>
<th>NN</th>
<th>Trainable Parameters</th>
<th>Learning Rate</th>
<th>Batch Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF – FC_{ECI}</td>
<td>23983</td>
<td>0.0005</td>
<td>25</td>
</tr>
<tr>
<td>FF – FC_{PBI}</td>
<td>270543</td>
<td>0.0010</td>
<td>25</td>
</tr>
<tr>
<td>CNN</td>
<td>63671</td>
<td>0.0005</td>
<td>25</td>
</tr>
<tr>
<td>DCNN</td>
<td>89073</td>
<td>0.0001</td>
<td>100</td>
</tr>
<tr>
<td>DINN</td>
<td>313071</td>
<td>0.0001</td>
<td>100</td>
</tr>
<tr>
<td>IRNN</td>
<td>804804</td>
<td>0.0001</td>
<td>100</td>
</tr>
</tbody>
</table>

[Ref 4] Blatzheim, Böckenhoff et. al. submitted to NF 2019
(\(\nu, \Delta R\)) - Reconstruction

Relative error < 3%

Relative error < 2%

[Ref 4] Blatzheim, Böckenhoff et. al. submitted to NF 2019
Conclusion

• Achievements:
  • Successfull reconstruction of rotational transform and radial axis shift \((\iota, \Delta R)\) from heat load images
  • Simulation speedup (factor 20)
  • CNNs and further image algorithms applicable
  • Compatible for synthetic and experimental data
  ➢ Improve performance of sparse experimental data!
  • Identification of critical states (overload)

Outlook

• Next steps:
  • Application to experimental divertor data
  • Virtual RL control test
APPENDIX
IR Data pre-processing

Experiment → Dias IR camera

Mapping → Map to CAD using known geometry

Heat-Flux → Solving $\Gamma(t)$ with 1D code THEODOR

Interpolation → Interpolate mesh between pixels

[Fabio Pisano] [Holger Nieman]
Limiter and Divertor Infra-Red (IR) Observation
Conclusion

- Successfull reconstruction of rotational transform and radial axis shift ($\iota$, $\Delta R$) from heat load images
Real Time Heat Load Control: Reinforcement Learning

- **Agent:**
  - Tow independent NNs determine
    - \( Q_\pi(s, a) = r(s_t, a_t) + \gamma \max_{a'} Q_\pi(s', a') \)
  - Loss function for both:
    - \( L = Q_{\pi,1}(s, a) - Q_{\pi,2}(s, a) \)
    - \( a_t = \max_a Q_\pi(s_t) \)

- **States could be**
• **OP 1:**
  - No significant variance in $\beta$
  - $\iota$-scan with 6 configurations only (multiple experiments and frames, “config. $J$”)

• **Curse of dimensionality**
  $\begin{align*}
  (I_1, \ldots, 5, A_B, S_1, S_2, n, T, I_{tor}, E_r, \ldots)
  \end{align*}$
  - Sparse experimental data
  - Clustered data $\Rightarrow$ Inter- / Extrapolation
Data Set(s)

• $\tau$ Scan (“Config. J“ – Op1.1)

<table>
<thead>
<tr>
<th>Infrared</th>
<th>Synthetic</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Data Set Size

$|\mathbb{I}| = 319$

$|\mathbb{S}| = 3993$
Difference in Experimental and Synthetic Data

Synthetic data has broader strike line.

Special Artifacts (e.g. hot spots) can not be resolved but will not appear in simulation.
**Strikeline Complexity: Divertor vs Limiter**

Plasma – Geometry interaction is less complex for limiter than divertor.

- Divertor Strike Line has more features
- NN performance better on Divertor than on Limiter Strikeline
Heat Load on Limiter

Böckenhoff et al. 2018

Simulation
~ 4000

Experiment
~ 300
Reconstruction of $I_B$ with NN: IR Data

(Ⅱ, Ⅱ, Ⅱ)

Features
- Reconstruction works
- Significant Scattering
- (too) Few Experiments in this iota scan
Reconstruction of $I_B$ with NN: IR Data

$(\bar{I}, \bar{I}, \bar{I})$
Reconstruction of $I_B$ with NN: IR Data

$(\Pi, \Pi, \Pi)$

Lessons

- Significant Scattering
- (too) Few Experiments
Reconstruction of $I_B$ with NN: Synthetic Data

$\left( S, S, S \right)$

**Lessons**
- Good Performance
- Only relevant in experimental context
Simulation Trained, Experiment-Tested

$\left( S, S, I \right)$
NN Input Processing

Partitioning

**Parametrization**
- Center of Mass & Standard Deviation
- Center of Mass & First Main Axis
- Ratio to Maximum Heat Load

2 x 2
Training Mixed Different Parametrizations

(M, M, I)

M = SUI
Training Mixed Different Parametrizations

\((M, M, I)\)

Reconstruction Quality

\(\text{rmse}(M_{90}, M_{10}, I_{C})(\rho)\)
\(\text{rmse}(M_{90}, M_{10}, I_{C})(\mu, \sigma)\)
\(\text{rmse}(M_{90}, M_{10}, I_{C})(\mu, \delta)\)

10\% of \(I_B\) range

Partitioning Resolution:

2 x 1, 4 x 1, 9 x 1, 9 x 5, 15 x 6, 18 x 8, 27 x 10, 36 x 12, 72 x 15, 144 x 30
Training Mixed Different Parametrizations

$\left( M, M, I \right)$

$\text{rmse}(M_{90}, M_{10}, I_C)(\rho)$
$\text{rmse}(M_{90}, M_{10}, I_C)(\mu, \sigma)$
$\text{rmse}(M_{90}, M_{10}, I_C)(\mu, \delta)$

10% of $I_B$ range

Resolution

$2 \times 1$, $4 \times 1$, $9 \times 1$, $9 \times 5$, $15 \times 6$, $18 \times 8$, $27 \times 10$, $36 \times 12$, $72 \times 15$, $144 \times 30$
Training Mixed Different Parametrizations

$\mathbf{M}, \mathbf{M}, \mathbf{I}$

**Graph Details:**
- **x-axis:** Resolution
- **y-axis:** $\text{rmse}$
- **Legend:**
  - $\text{rmse}(M_{90}, M_{10}, I_C)$ ($\rho$)
  - $\text{rmse}(M_{90}, M_{10}, I_C)$ ($\mu, \sigma$)
  - $\text{rmse}(M_{90}, M_{10}, I_C)$ ($\mu, \delta$)
  - 10% of $I_B$ range

**Graph Notes:**
- The graph shows the $\text{rmse}$ for different parametrizations across various resolutions.
- The $x$-axis represents different resolution sizes.
- The $y$-axis represents the $\text{rmse}$ values, with a logarithmic scale.
- The legend provides details on the different parametrization sets evaluated.
NN learns to only extract relevant features from the simulation.
Representative Performance

\[(M, M, I)\] Median of \( \text{rmse}(M_{90}, M_{10}, I_{C})_{P}^{18 \times 8, \text{Conv}} \)
Neural Networks (Limiter)

Blatzheim et al. 2018
Generative Adversarial Networks

Neural Networks

G: Generator
D: Discriminator

Loss small, if discriminator correct

Loss small, if discriminator incorrect
Basic Regression Approach

Neural Networks
R: Regression
Plasma GAN

Neural Networks
G: Generator
D: Discriminator
R: Regression

Loss small, if discriminator correct
Loss small, if discriminator incorrect
Neural Networks (Divertor)
Neural Networks (Divertor)

Inception ResNet

Experiments with $i = 0$ and $i = 3$
Reinforcement Learning

Reinforcement learning

Interpreter → Agent → Environment

State → Reward → Action

Observation → Reinforcement learning

→ Computer → Gear

24.05.2019

Böckenhoff - Third Technical Meeting on Fusion Data Processing