13th IAEA Technical Meeting on Plasma Control Systems, Data Management and Remote Experiments in Fusion Research

IAEA

Applying machine learning in nuclear fusion data:

- **Reinforcement learning for building nuclear fusion classifiers from scratch**
- **Deep learning models to generate realistic new data in nuclear fusion**

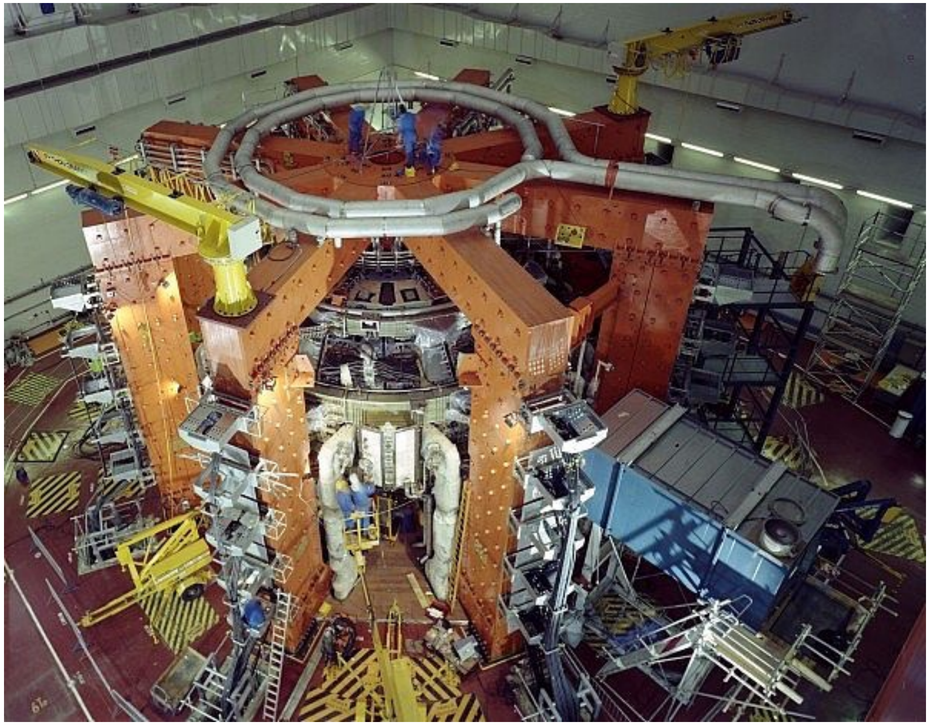
Gonzalo Farias[1]
Ernesto Fabregas[2]
Sebastián Dormido-Canto[2]
Jesús Vega[3]
Francisco Esquembre[4]
Ignacio Pastor[3]
Diego Hidalgo[1]
Sara Cuellar[1]
Heilym Ramirez[1]
Ricardo Correa[1]

July 8/2021, Vienna, AUSTRIA

1 PONTIFICIA UNIVERSIDAD CATOLICA DE VALPARAISO
2 UNED
3 GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
4 UNIVERSIDAD DE MURCIA

Introduction

- Experimental nuclear fusion devices have enormous databases



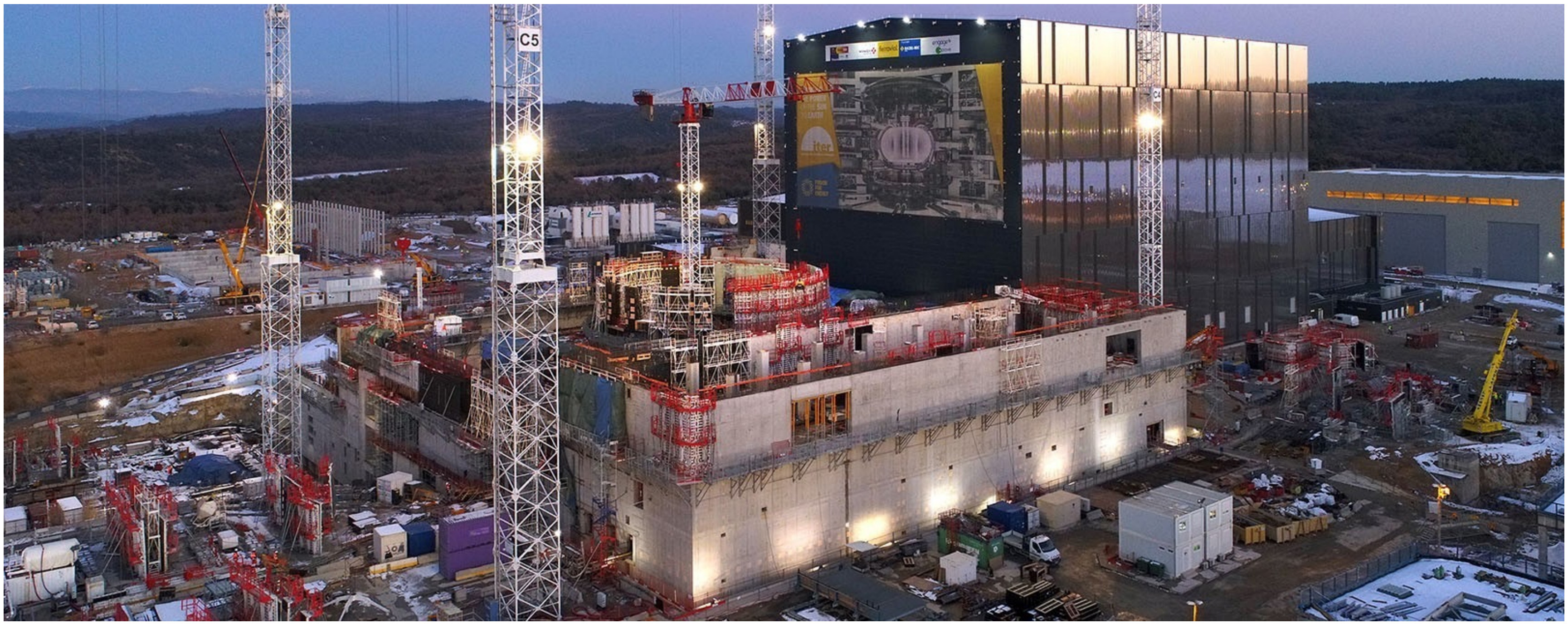
A simple shot of few seconds can generate huge quantity of data:

- **TJ-II** device has +1000 channels of measurements.
- A discharge in **JET** can take a couple of seconds (**10 GB/shot**. around 100 TB/year).

It is estimated that only 10% of this data is analyzed!

Introduction

- Experimental nuclear fusion devices have enormous databases



ITER could generate **1 TB/shot**. around 1 PB/year.

Machine learning algorithms requires data to work, but we can not wait until collect all the data to, for example, create disruption predictors!

Introduction

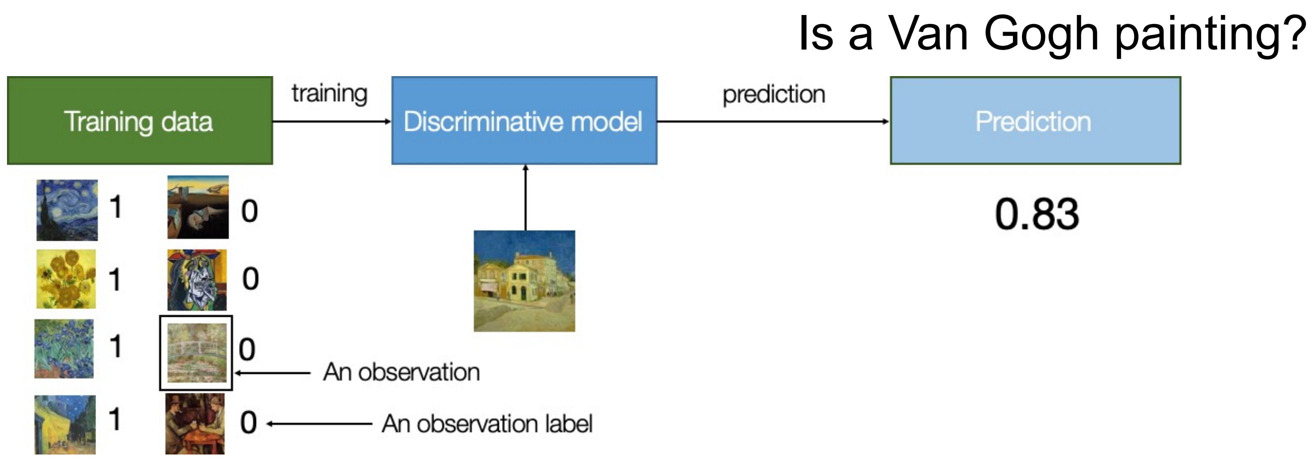
- Learning from "few" samples could involve:
  - Train predictors/classifiers **from scratch under data scarce conditions**. A reinforcement learning approach. **Discriminative modeling**.
    - Reinforcement learning (RL) uses a **reward function** to build the discriminative model (classifier). RL can be used to weight different features to classify the input data. The more data feed the model the more accuracy we have.

    Vega, J., Murari, A., Dormido-Canto, S., Moreno, R., Pereira, A., Acero, A., & JET-EFDA Contributors. (2014). **Adaptive high learning rate probabilistic disruption predictors from scratch for the next generation of tokamaks**. Nuclear Fusion, 54(12), 123001.

    Dormido-Canto, S., Vega, J., Ramírez, J. M., Murari, A., Moreno, R., López, J. M., ... & JET-EFDA Contributors. (2013). **Development of an efficient real-time disruption predictor from scratch on JET and implications for ITER**. Nuclear Fusion, 53(11), 113001.
  - Generate new samples following the training data distribution. Deep learning models to generate realistic new data. **Generative modeling**.

Background

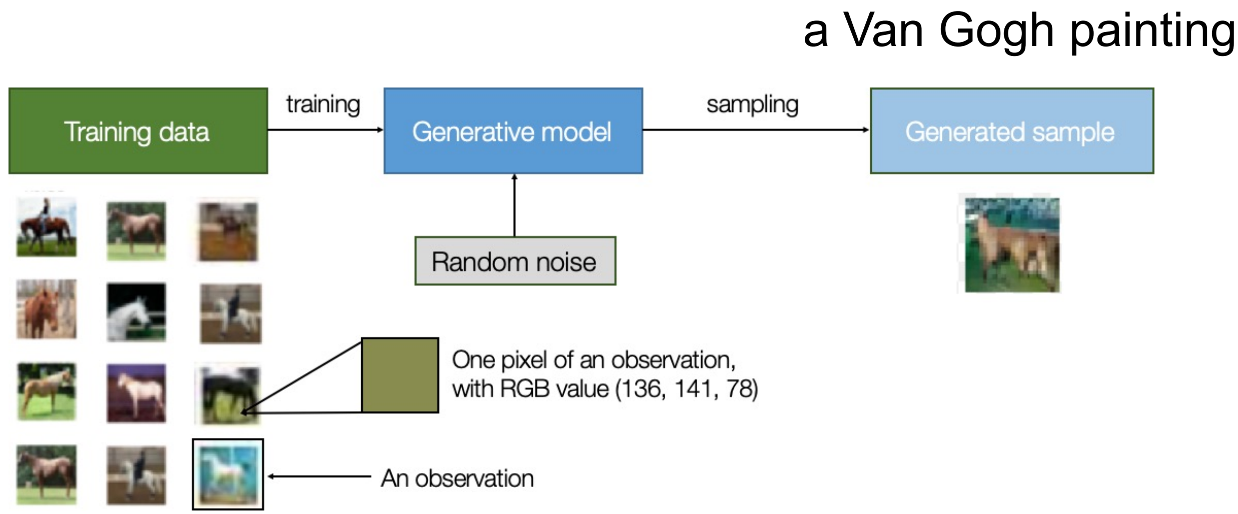
❑ Discriminative modeling



- A well-known model to predict/classify an observation (painting)
- The training data is used to learn a model (discriminative model)
- Each sample/observation has a label (1/0 means yes/no)
- Even if we were able to build a perfect discriminative model, it would still have no idea how to create an observation (painting)

Background

❑ Generative modeling



- A generative model mimics the unknown probabilistic distribution that explains each observation (painting)
- The training data is used to learn a model (generative model)
- No label is needed (unsupervised learning)
- The model generates new and distinct observations that look as if they have been included in the original training set

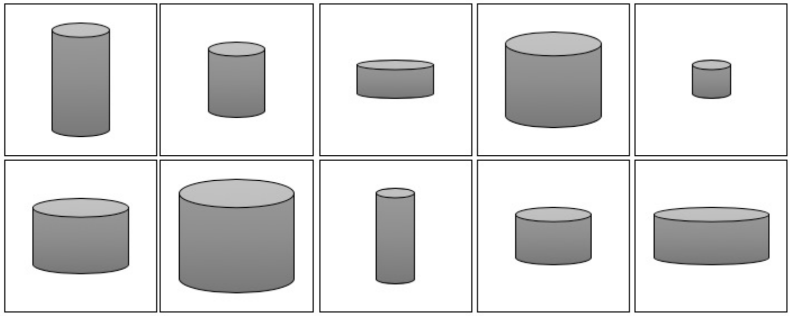
Background

❑ Generative modeling framework

- We have a dataset of observations X
- We assume that the observations have been generated according to some **unknown probabilistic distribution**
- A generative model tries to **mimic the unknown distribution**. If we success, we can sample from the model to generate new observations
- A model successes when:
  - It can generate examples **that seems to have the same unknown distribution** that the training dataset
  - It can generate **examples that are suitable different from the observations** in the training dataset

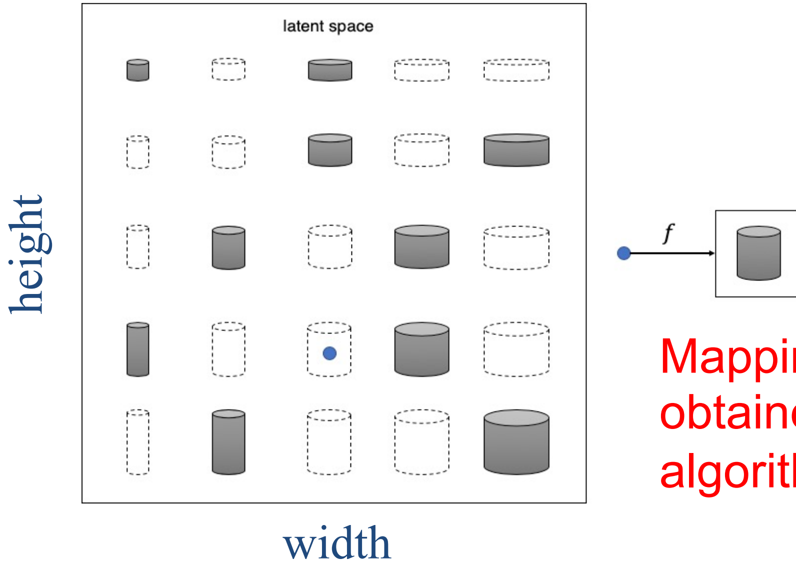
Generative modeling: Latent space

Training dataset (original space)



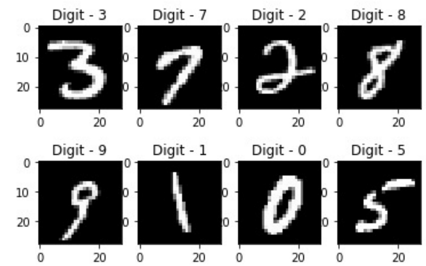
Instead of trying to model high-dimensional space directly, we should describe each observation in the traingin set using a low-dimensional **latent space** and learn a mapping function (***f***) to go from latent to original space (generating a new observation).

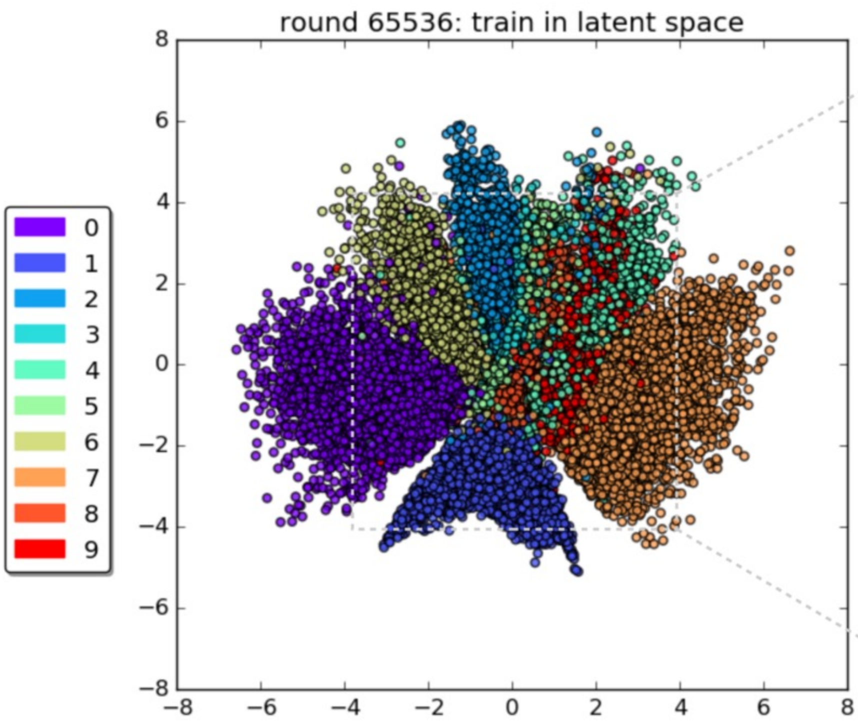
Latent space (low-dimensional space)

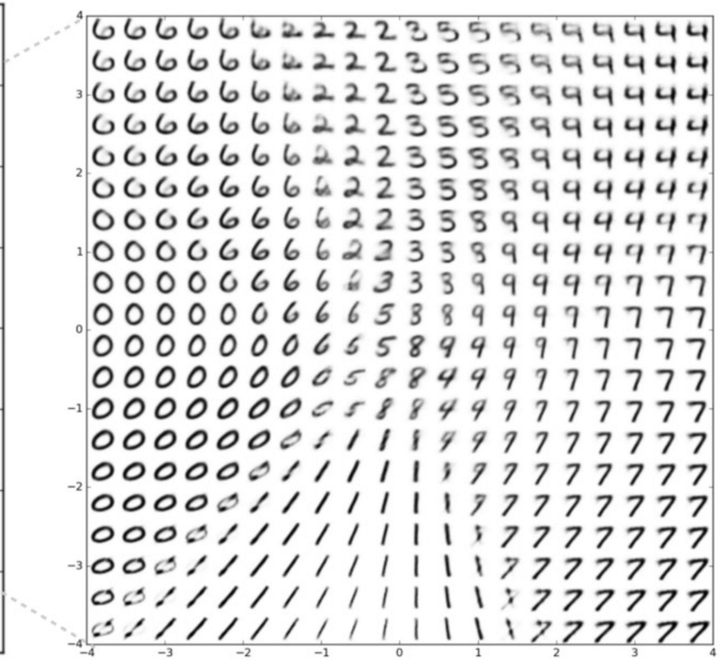


Mapping function ***f*** can be obtained with ML/DL algorithms

Generative modeling: Latent space

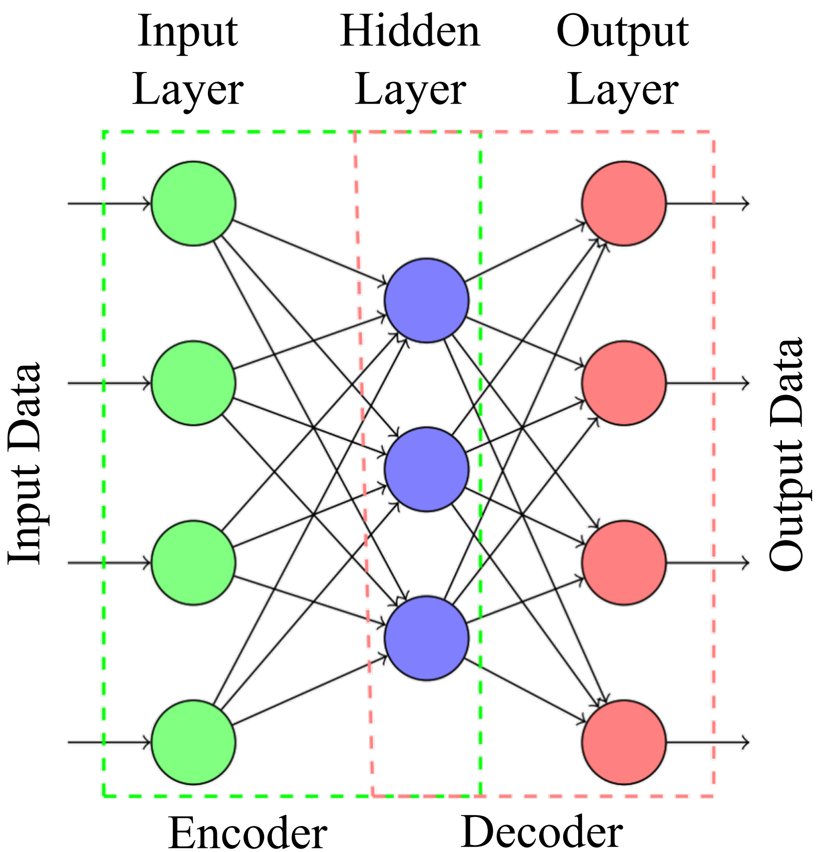






Generative modeling: Autoencoder

- Autoencoder (AE)



It adjusts the bias and weights to learn a function in an unsupervised way.

AE tries to learn an approximation to the identity function, so it outputs something similar to its input.

Placing constraints on the network, such as by limiting the number of hidden units, we can discover interesting structure about the data.

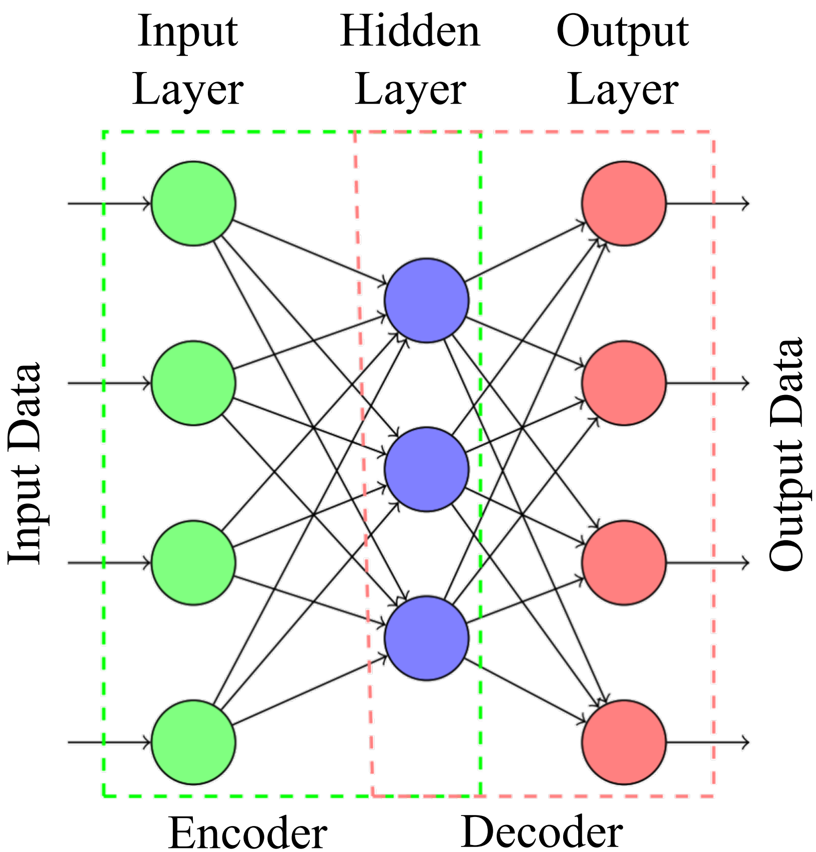
The network is forced to learn a "compressed" representation of the input.

Farias, G., et al. (2016). **Automatic feature extraction in large fusion databases by using deep learning approach**. Fusion Engineering and Design, Volume 112, Pages 979–983.

Farias, G., et al. (2018). **Applying deep learning for improving image classification in Nuclear Fusion Devices**. IEEE Access, vol. 6, pp. 72345–72356.

Generative modeling: Autoencoder

- Autoencoder (AE)



Cost Function

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(x_{kn} - \hat{x}_{kn})^2$$

$$\Omega_{weithts} = \frac{1}{2}\sum_{l}^{L}\sum_{j}^{n}\sum_{i}^{k}\left(w_{ji}^{(l)}\right)$$

$$\Omega_{sparsity} = \sum_{i=1}^{s_2}\rho \log\frac{\rho}{\hat{\rho}_i} + (1-\rho)\log\frac{1-\rho}{1-\hat{\rho}_i}$$

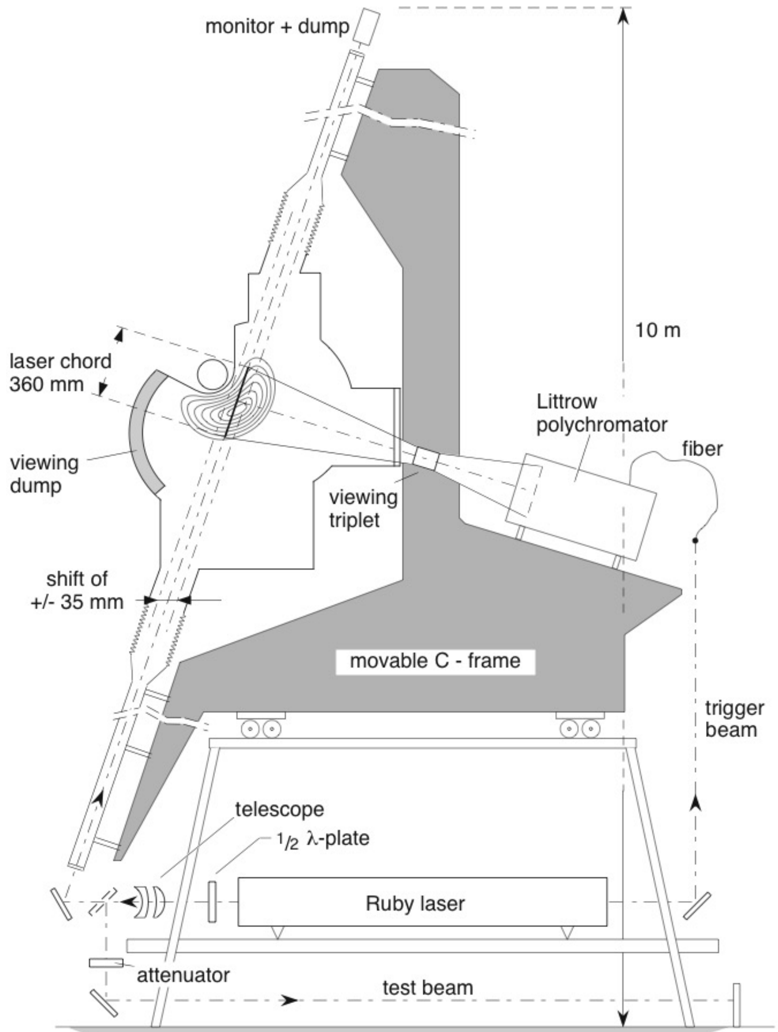
$$\hat{\rho}_i = \frac{1}{n}\sum_{j=1}^{n} h\left(w_i^{(l)T}x_j + b_i^{(l)}\right)$$

$$J = MSE + \lambda * \Omega_{weithts} + \beta * \Omega_{sparsity}$$

Farias, G., et al. (2016). **Automatic feature extraction in large fusion databases by using deep learning approach**. Fusion Engineering and Design, Volume 112, Pages 979–983.

Farias, G., et al. (2018). **Applying deep learning for improving image classification in Nuclear Fusion Devices**. IEEE Access, vol. 6, pp. 72345–72356.

TJ-II Thomson Scattering diagnostic

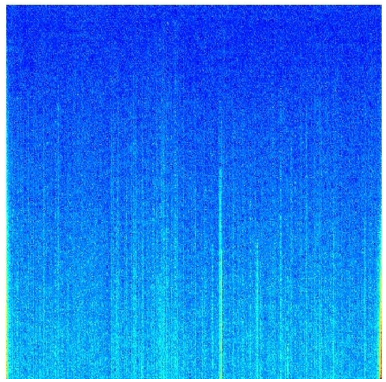


The **Thomson Scattering (TS) diagnostic** of the TJ-II stellarator provides **temperature and density profiles**.

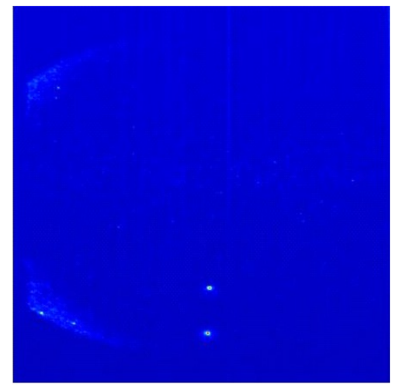
The diagnostic acquires **five types of images** (spectra of laser light scattered by plasma): CCD camera background (**BKG**), measurement of stray light without plasma (**STR**), during electron cyclotron resonant heating (**ECH**), during neutral beam injection (**NBI**), and after reaching the cut-off density during electron cyclotron resonant heating (**COF**).

TJ-II Thomson Scattering diagnostic

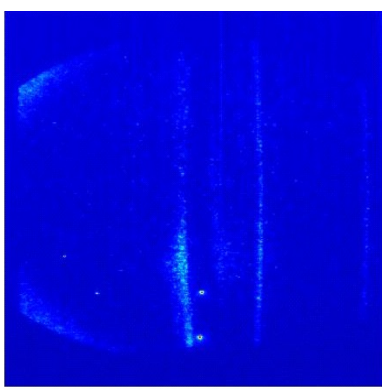
BKG



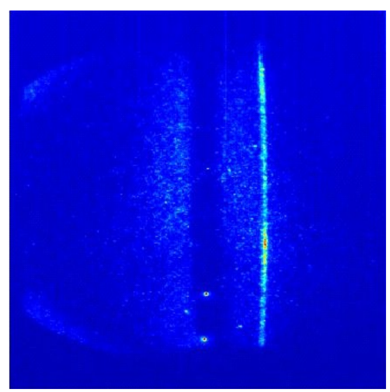
STR



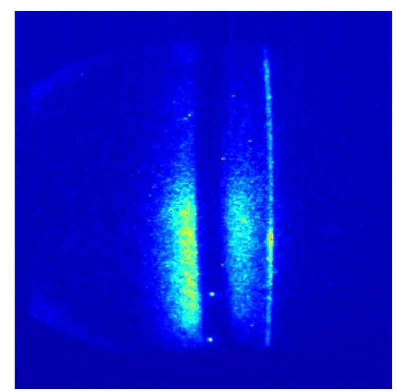
ECH



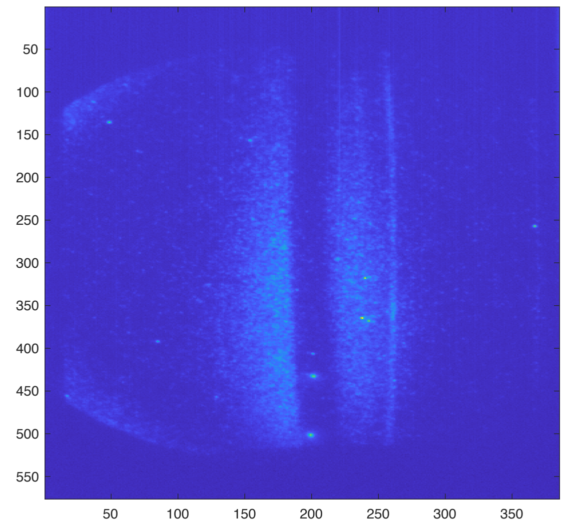
NBI



COF

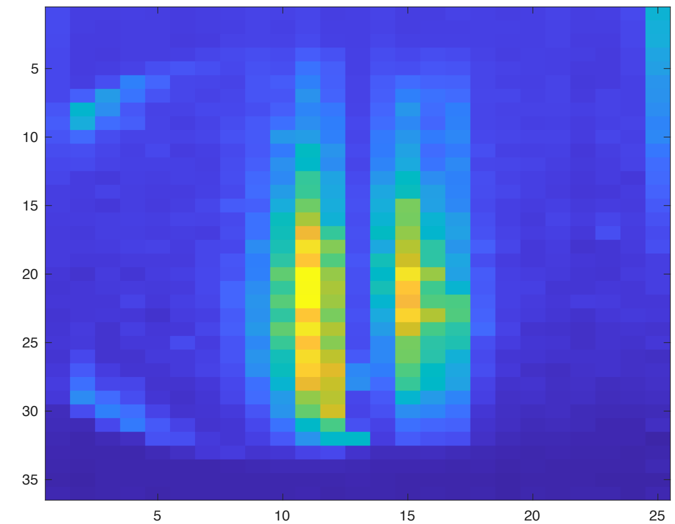


TJ-II Thomson Scattering diagnostic



Wavelet transform

Feature reduction

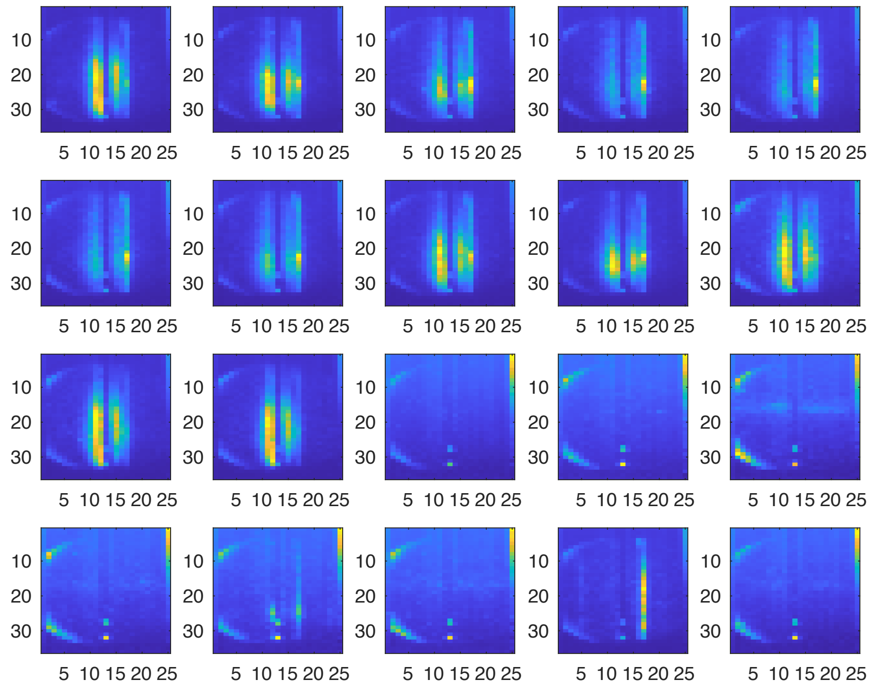


Original dimensionality
576x385 pixels

Reduced dimensionality
36x25 pixels

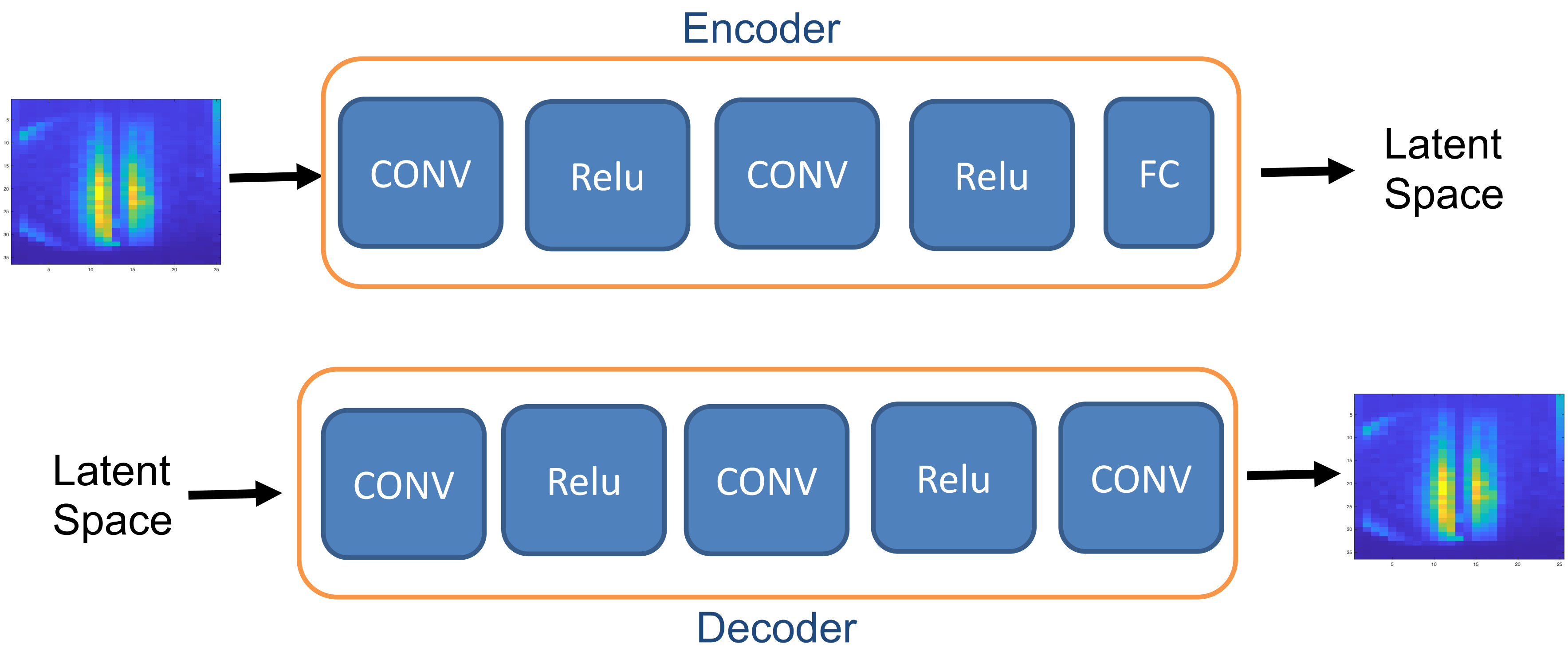
TJ-II Thomson Scattering diagnostic

training dataset (242 samples)



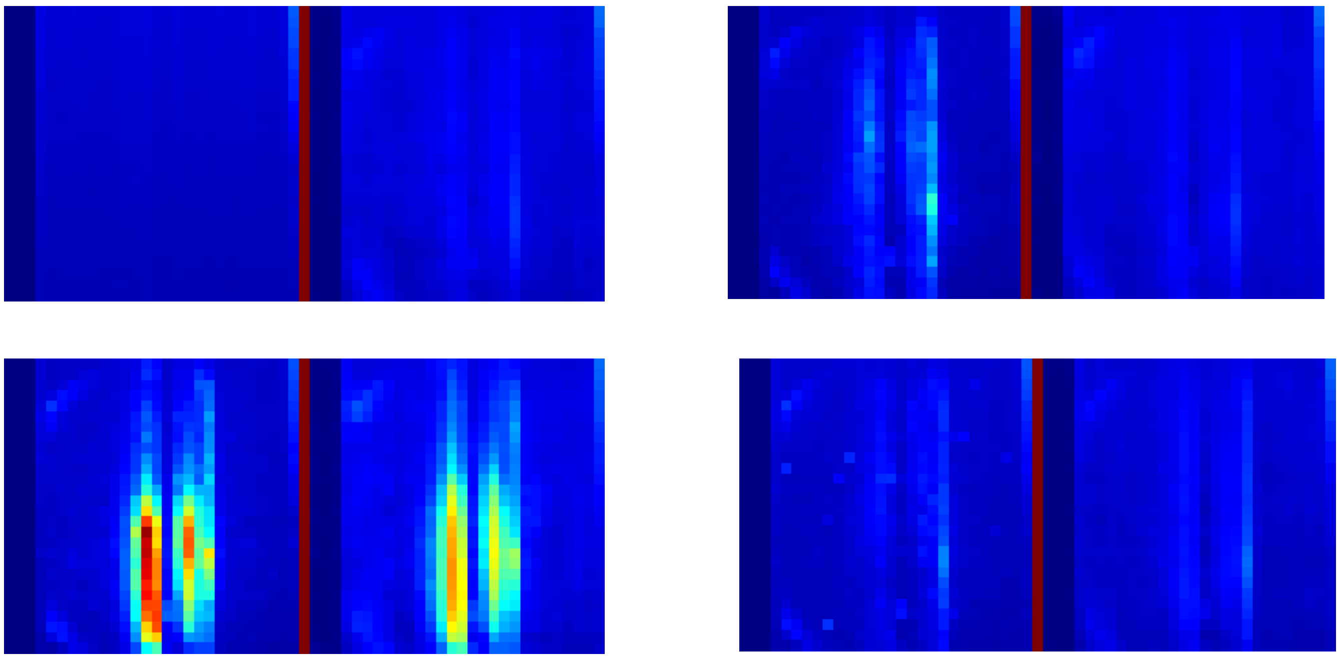
TJ-II Thomson Scattering diagnostic

Autoencoder (variational)



TJ-II Thomson Scattering diagnostic

Some examples of reconstruction

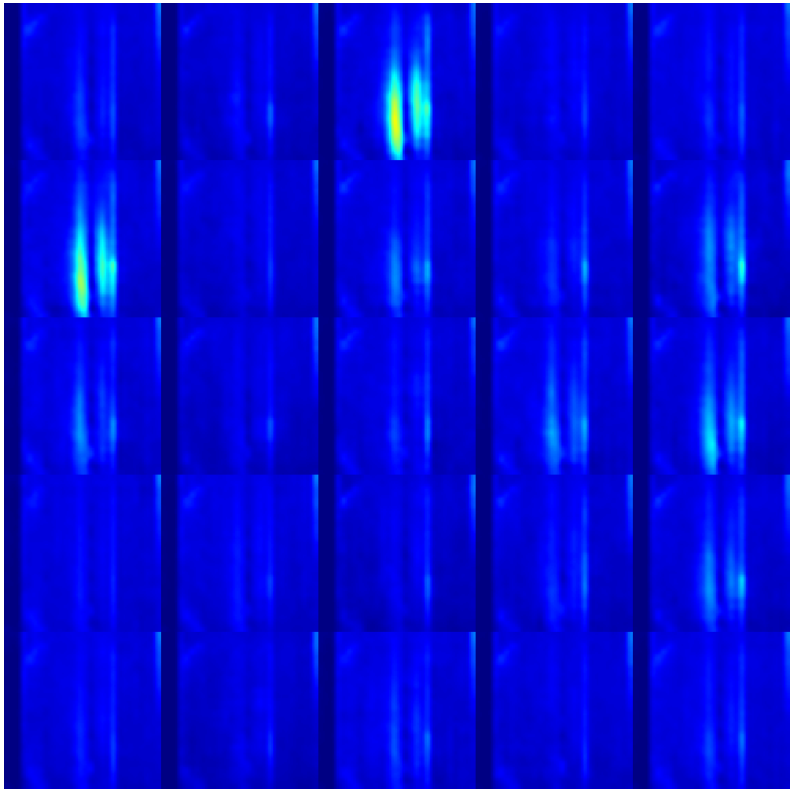


Original Reconstruted Original Reconstruted

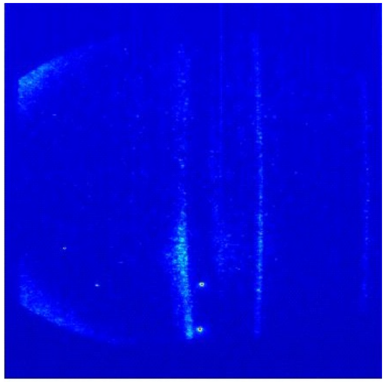
TJ-II Thomson Scattering diagnostic

Some examples of generated TS images

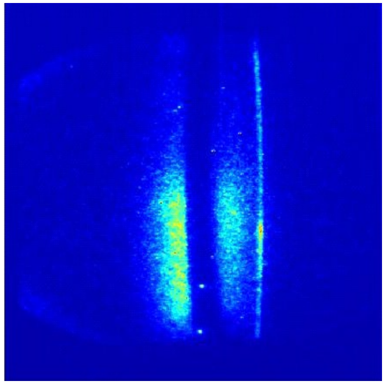
Generated samples of TS images



ECH

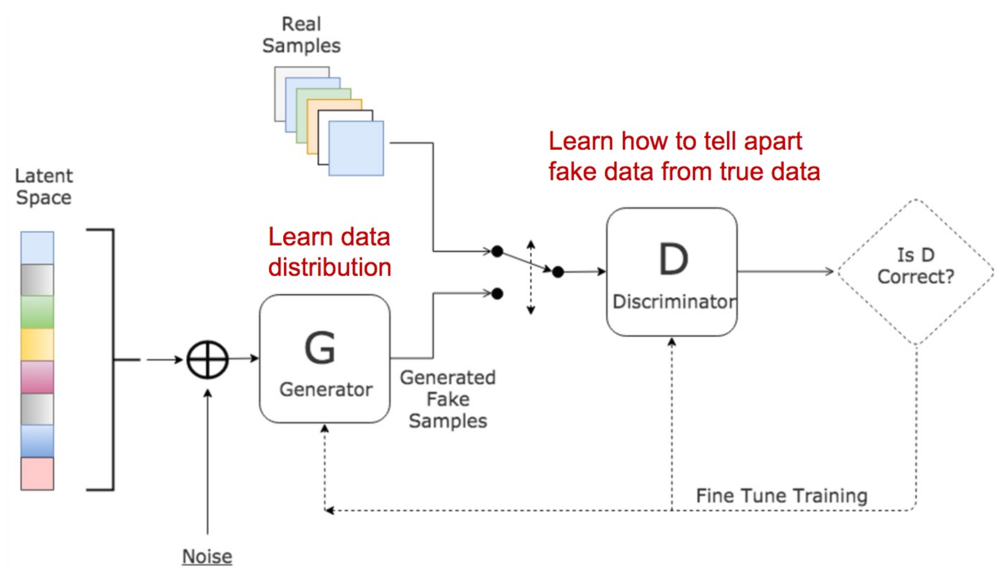


COF



Summary – Discussion

- ❑ Datasets are essential for machine learning algorithms
  - ❑ Under scarce conditions we can train models from scratch
  - ❑ Sometimes we could need to generate new and realistic data.
- ❑ AEs are a simple approach to generate new data.
- ❑ Generative adversarial networks (GANs)



13th IAEA Technical Meeting on Plasma Control Systems, Data Management and Remote Experiments in Fusion Research

IAEA

Applying machine learning in nuclear fusion data:

- **Reinforcement learning for building nuclear fusion classifiers from scratch**
- **Deep learning models to generate realistic new data in nuclear fusion**

Gonzalo Farias[1]
Ernesto Fabregas[2]
Sebastián Dormido-Canto[2]
Jesús Vega[3]
Francisco Esquembre[4]
Ignacio Pastor[3]
Diego Hidalgo[1]
Sara Cuellar[1]
Heilym Ramirez[1]
Ricardo Correa[1]

July 8/2021, Vienna, AUSTRIA

1 PONTIFICIA UNIVERSIDAD CATOLICA DE VALPARAISO
2 UNED
3 GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
4 UNIVERSIDAD DE MURCIA