Hyper-parametric search using HPC infrastructures for Tensorflow.

Gonzalo Ferro - CESGA.

Fortissimo experiment 707: Cyber-Physical Laser Metal Deposition (CyPLAM)

Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

Experiment Partners:

End User: EMO
Die and mould manufacturer (Slovenia)

Domain Expert & ISV: AIMEN
Expert in laser technologies (Vigo, Spain)

HPC Centre: CESGA
HPC expert and provider (Santiago de Compostela, Spain)

Co-funded by the European Union

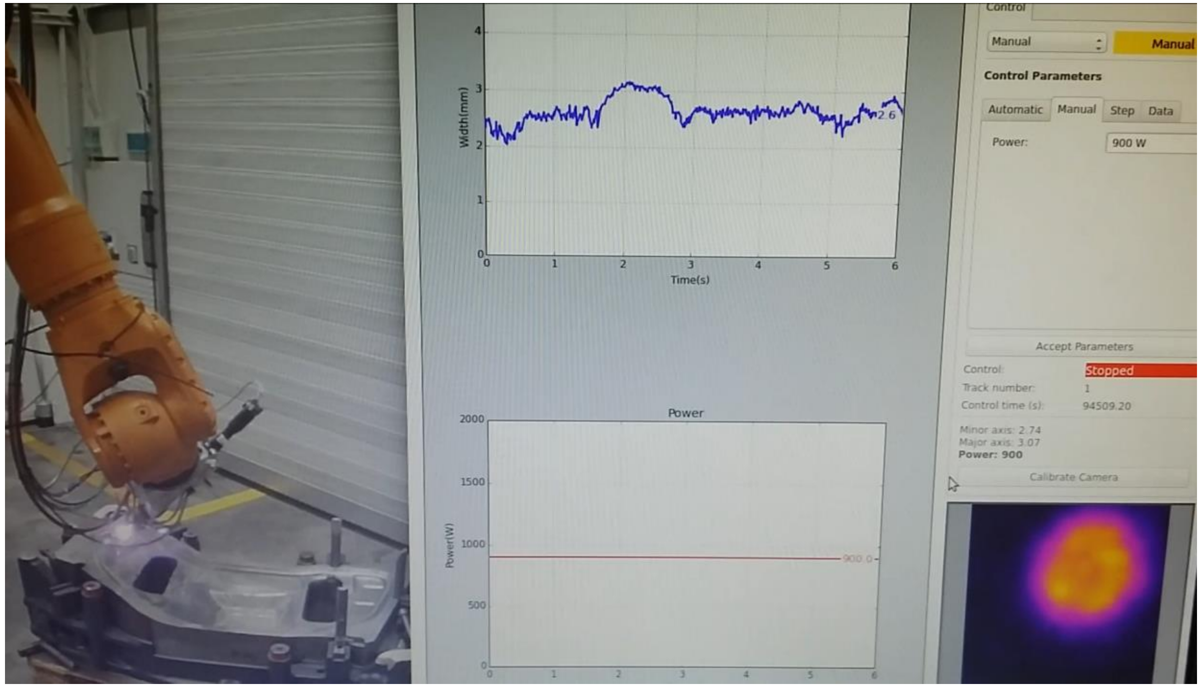
This project has received funding from the European Union’s H2020 research and innovation programme under grant agreement No 680481

Experiment Description:

Laser Metal Deposition (**LMD**) for building and repairing large metal parts.

LMD process recorder by Medium Wavelength Infrared (**MWIR**) sensors attached to laser header.

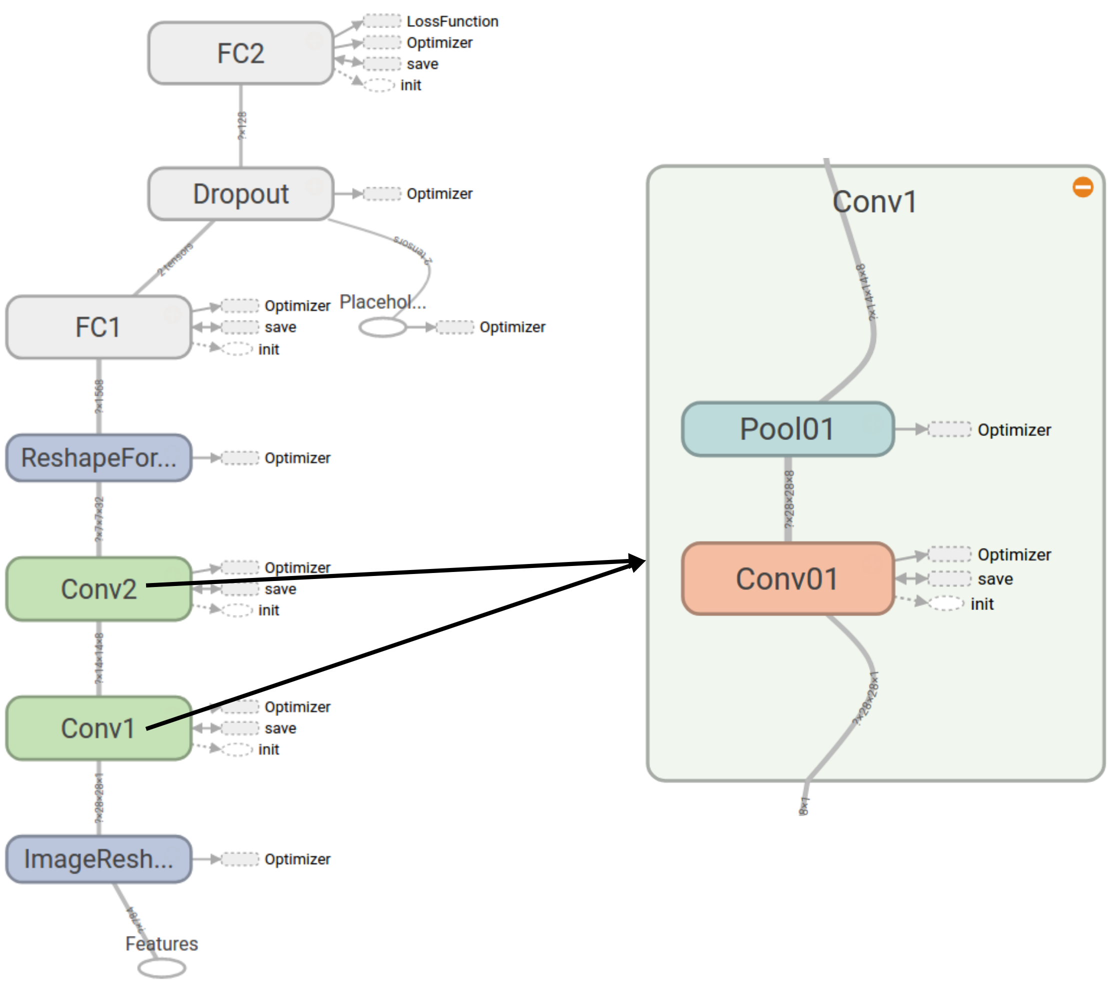
Use Convolutional Neural Network model for monitoring the LMD process based on the MWIR images.



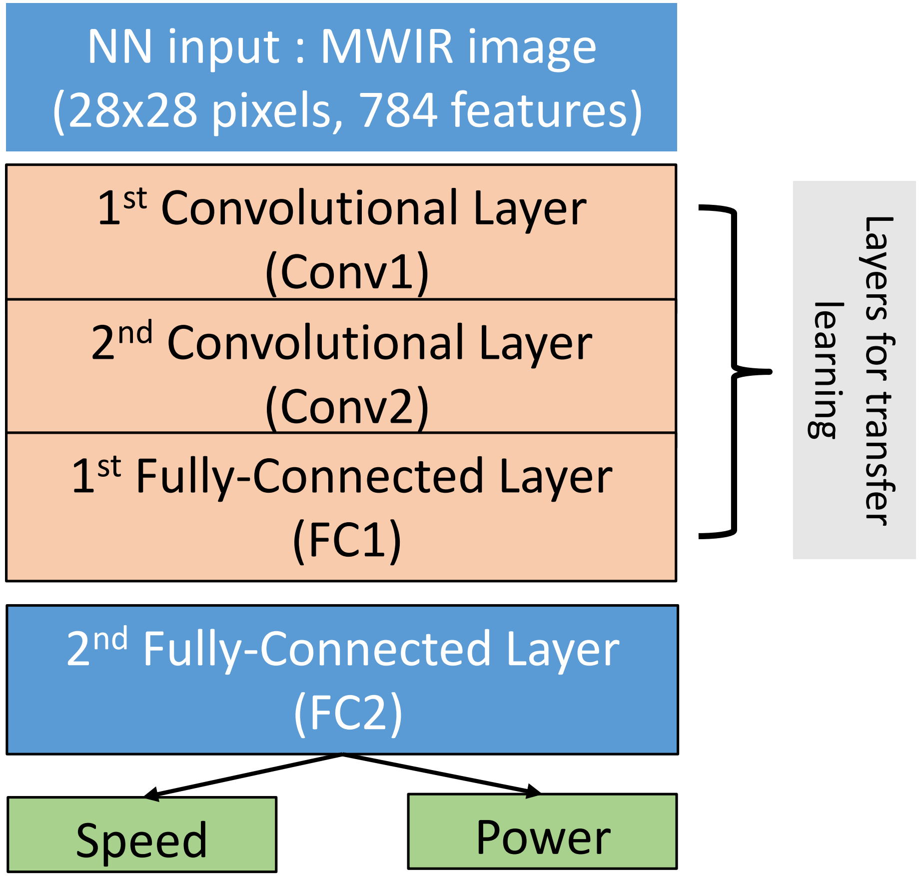
Co-funded by
the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

Model Description



NN Graph model (extracted from Tensorboard)



Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Experiment Description: Loss Function

$$Loss = \frac{1}{2m}\sum_{i=1}^{m}\left[\left(h_{LP}\left(x^i\right) - y_{LP}^i\right)^2 + \left(h_{LS}\left(x^i\right) - y_{LS}^i\right)^2\right]$$

- x^i is the ith frame of the training dataset.
- $h_{LP}\left(x^i\right)$ is the prediction for Laser Power of the NN model for ith frame.
- y_{LP}^i : label for Laser Power of the ith frame.
- $h_{LS}\left(x^i\right)$ is the prediction for Laser Speed of the NN model for ith frame.
- y_{LS}^i : label for Laser Speed of the ith frame.
- m : total number of training frames

Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Experiment Description: Metrics

$$Err^{i}(x^{i}, y_{LP}^{i}, y_{LS}^{i}) = Max\left(\frac{y_{LP}^{i}}{h_{LP}(x^{i})} - 1, \frac{y_{LS}^{i}}{h_{LS}(x^{i})} - 1\right)$$

$$Accuracy_{20\%} = \frac{1}{n}\sum_{i=1}^{n}\left(Err^{i}(x^{i}, y_{LP}^{i}, y_{LS}^{i}) < 20\%\right)$$

Co-funded by the European Union

This project has received funding from the European Union’s H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Experiment Description: Dataset

414 LMD tracks stainless steel (316L), powder & base material.

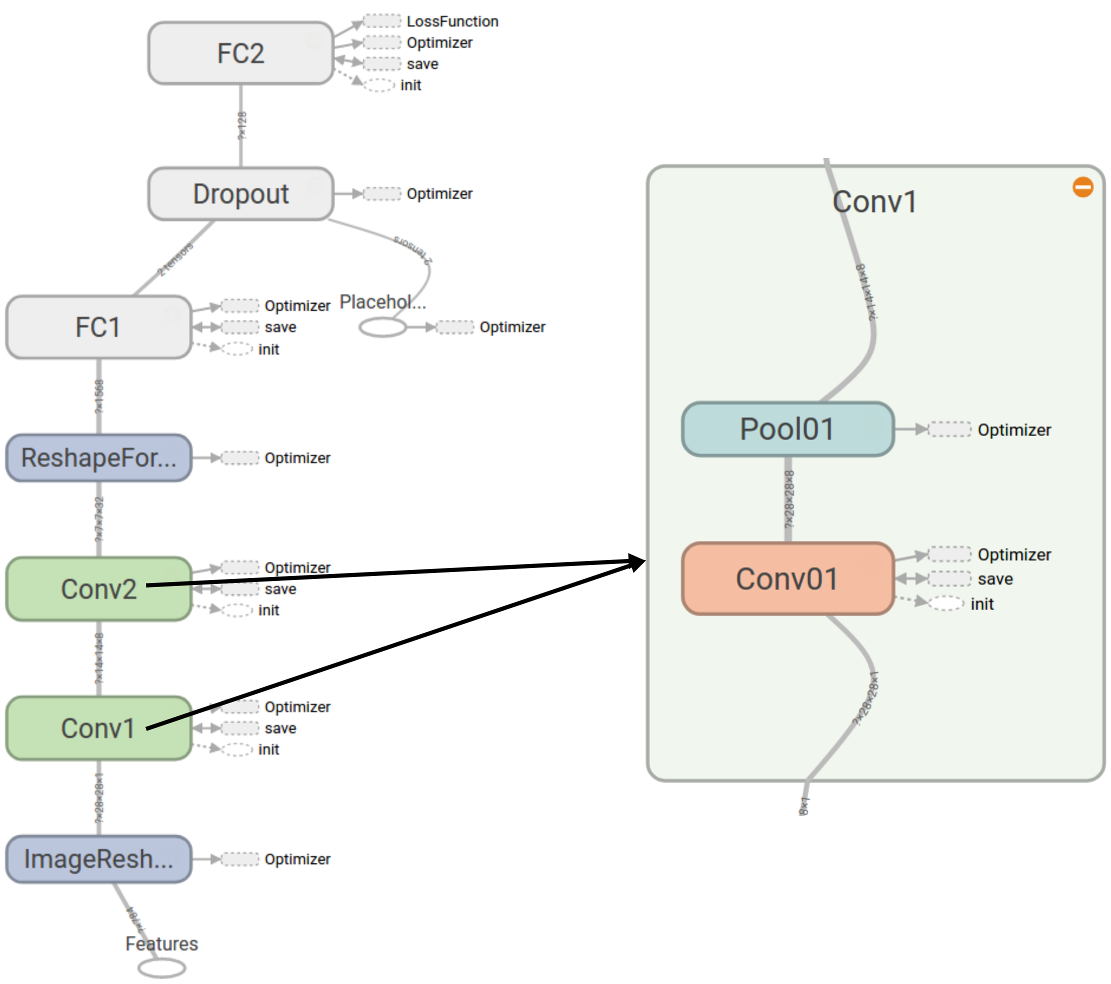
324 tracks for training | **90** tracks for validation

Dataset	Number of frames	Pixels	Number of labels
Training	44,505	28 x 28 = 784	2
Validation	162,216	28 x 28 = 784	2

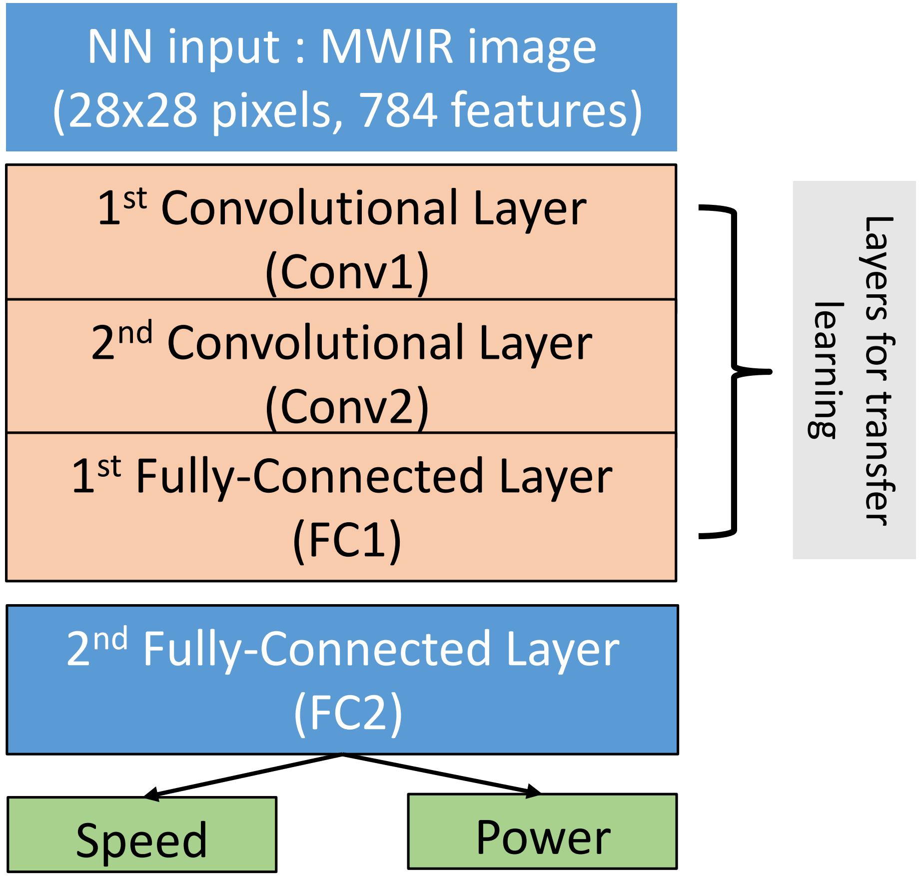
Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

Model Description



NN Graph model (extracted from Tensorboard)

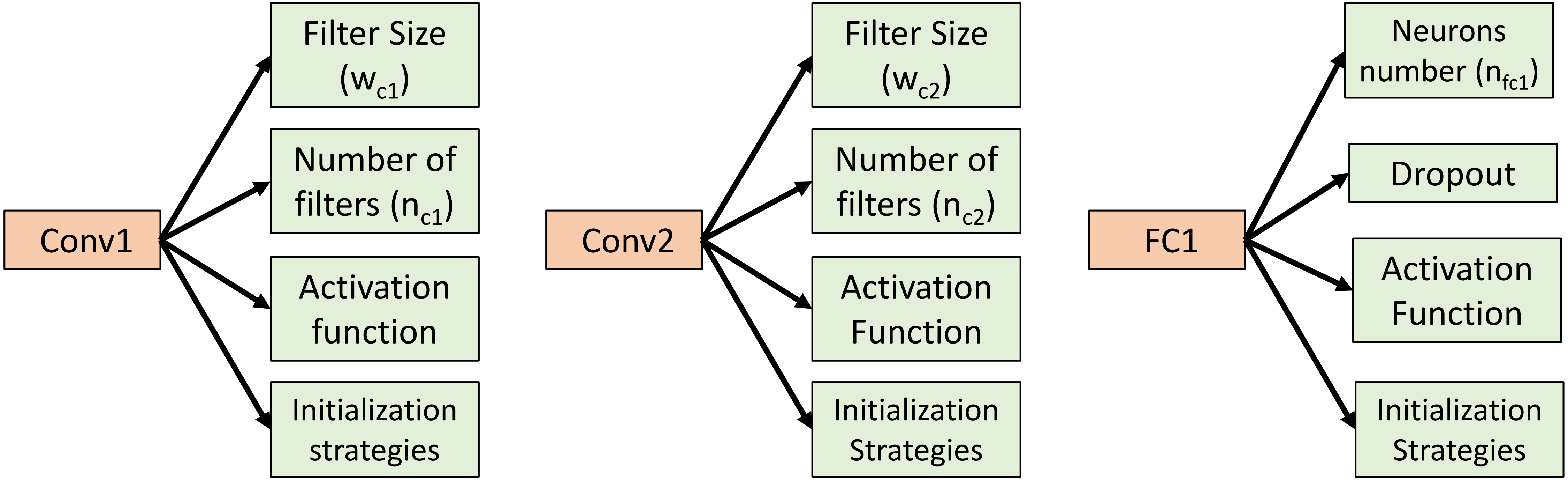


Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

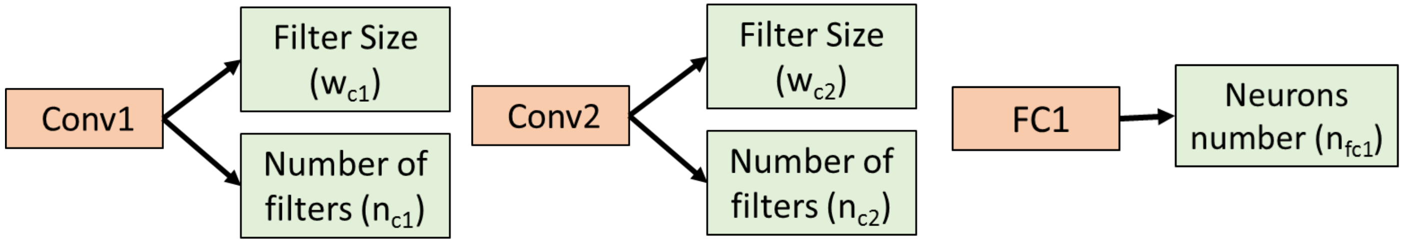
Hyper Parametric Search: Model Parameters



Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

Hyper Parametric Search: Implementation



	Filter Size	Number of Filters
Convolutional Layer 01	w_{c1}=3,5	n_{cv1}= 8, 16, 32
Convolutional Layer 02	w_{c2}=3,5	n_{cv2}= 32, 64
	Number of Neurons	
Fully-connected Layer	n_{fc1}=128, 256, 384, 512, 640, 768, 896, 1024	

Hyper Parameter	Value
Learning Rate:	10^{-5}
Dropout:	0.75
Batch size:	128
Optimizer:	Adam
Regularizers:	None
Iterations	400k

Grid search strategy

192 possible models!!!

200k-3M free parameters

Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Results: time performance

192 Models: 40 days

factor 13.7

70 hours

Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

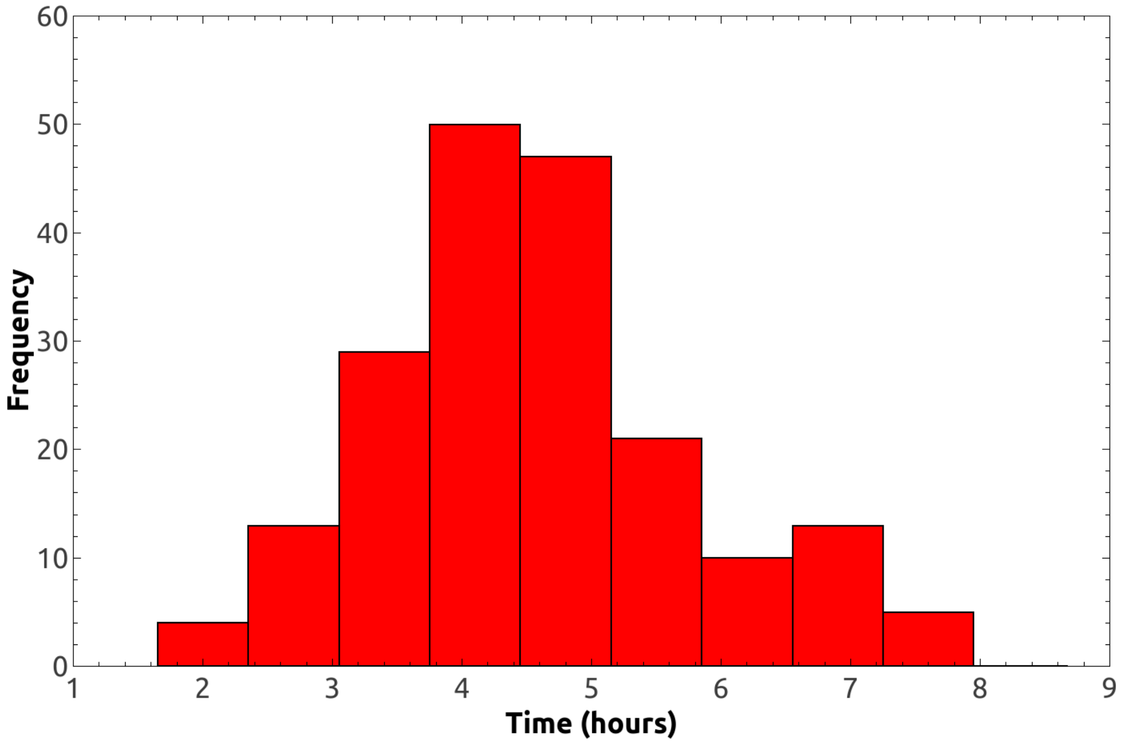
Results: time performance

192 Models: 40 days → 70 hours

Average trainings: 4-5 hours

Shortest training: 2.4 hours

Largest training: 8.2 hours



13.7 → 100 !!!!

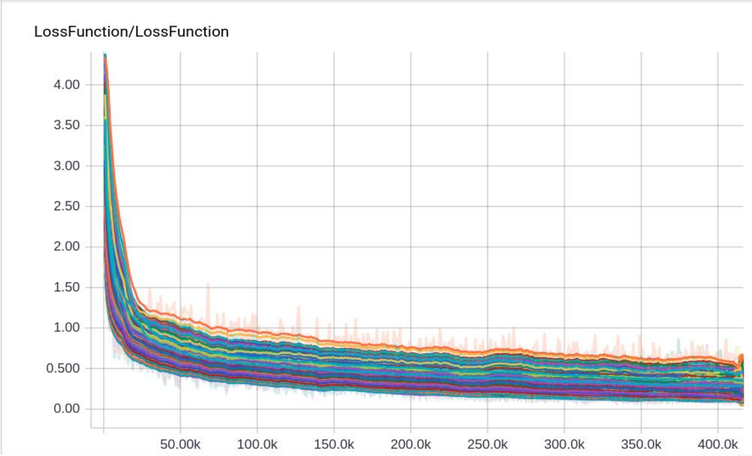
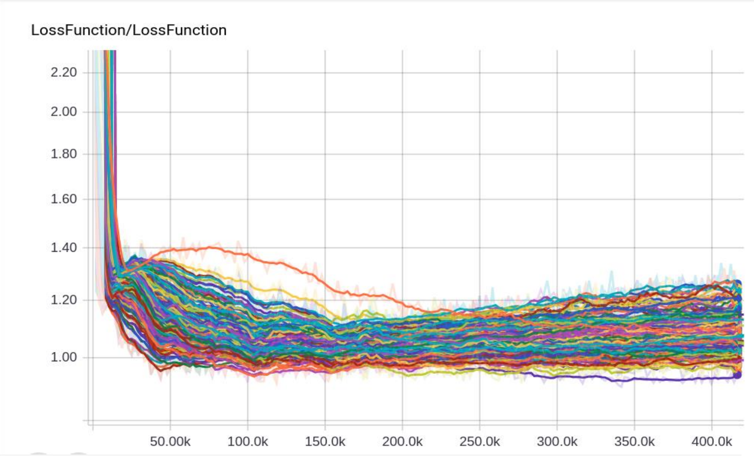
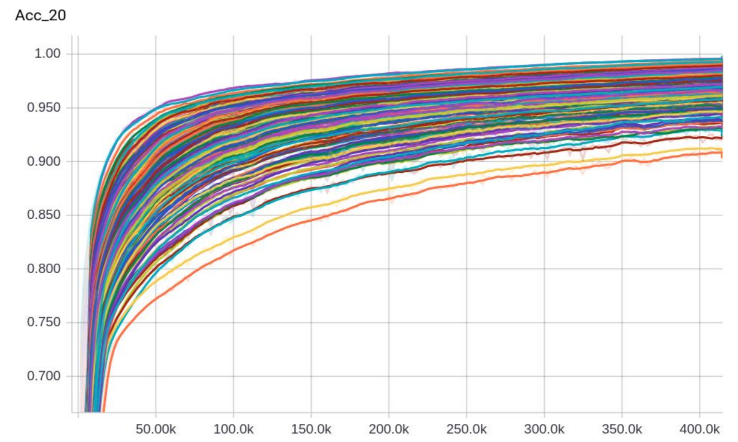
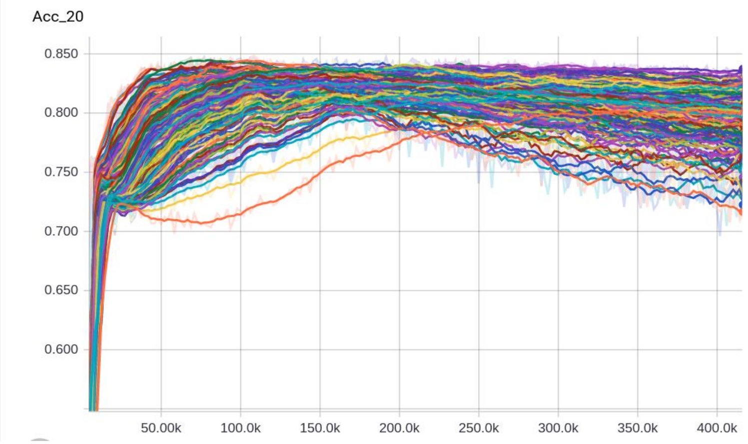
Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Results: Overall Performance

	Training dataset	Validation dataset
Loss function		
Accuracy$_{20\%}$		

Co-funded by
the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

Results: Overall Performance

FORTISSIMO

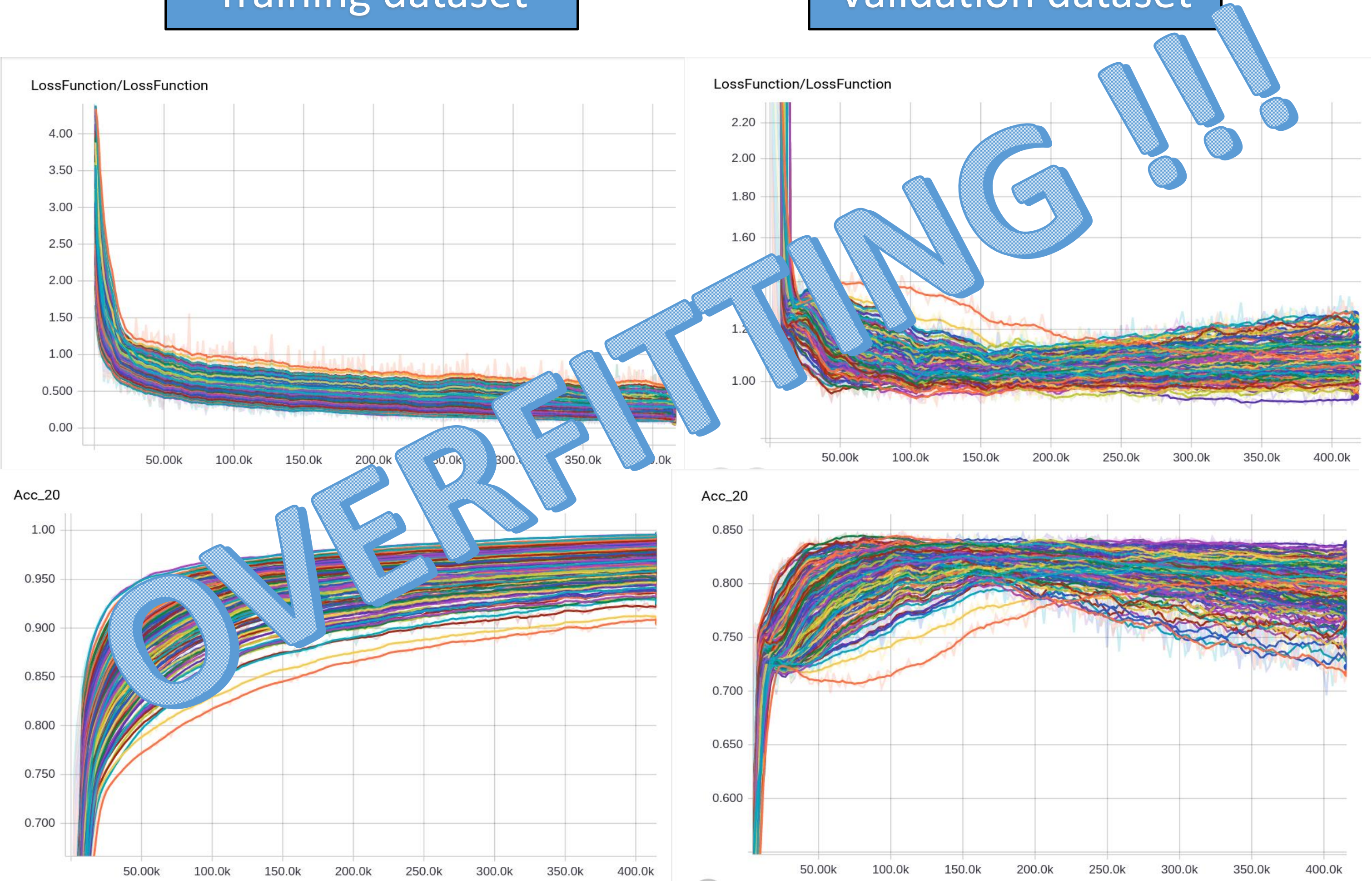
CESGA

Training dataset

Validation dataset

Loss function

Accuracy$_{20\%}$



Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Results: Overfitting solutions

- ✓ Increase size of training dataset.
- ✓ Change of activation functions.
- ✓ User of Regularization techniques.
- ✓ Early Stopping.

Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Results: Overfitting solutions

- ✓ Increase size of training dataset.
- ✓ Change of activation functions.
- ✓ User of Regularization techniques.
- ✓ Early Stopping.

EARLY STOPPING !!!

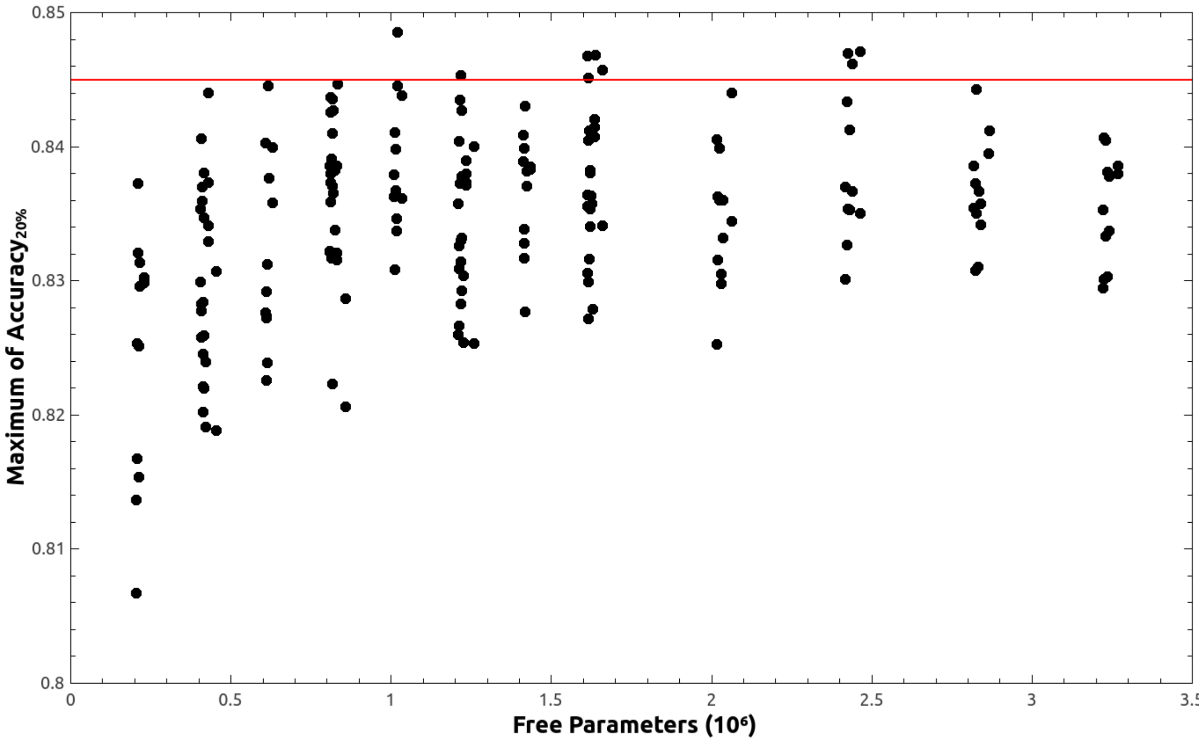
Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Results: Early Stopping



Top 10 models

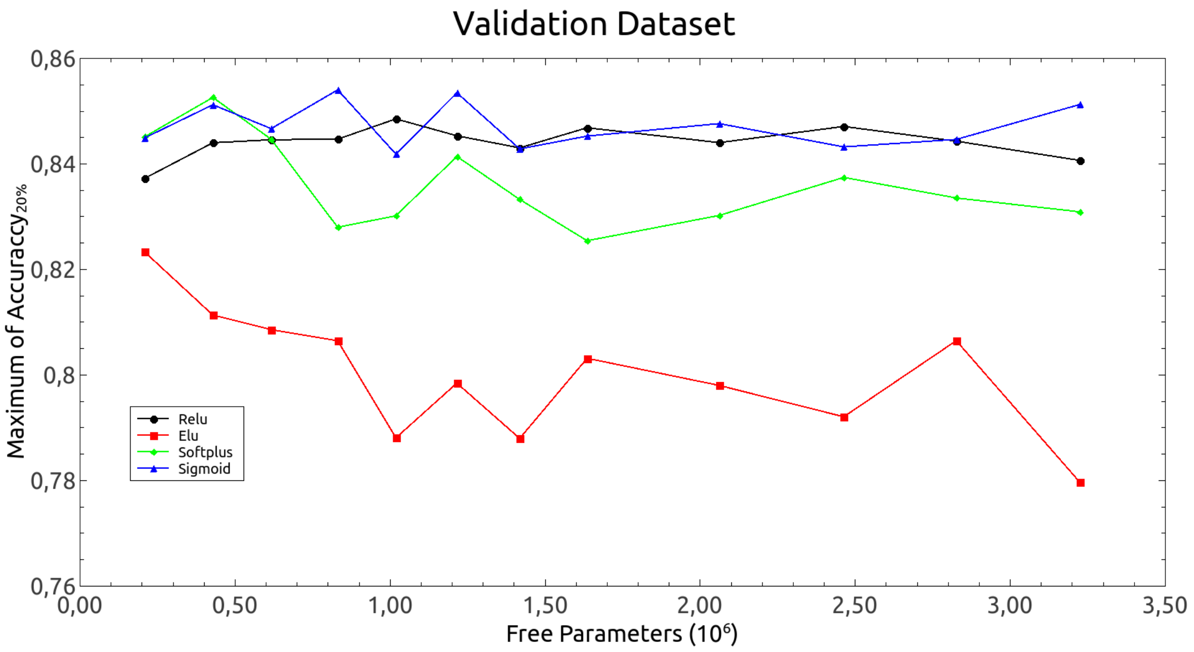
	w_{c1}	n_{c1}	w_{c2}	n_{c2}	n_{fc1}	Max Accuracy$_{20\%}$	Iteration Step	Free Parameters
Model-108	5	16	5	32	640	0.8485	108,264	1,018,690
Model-189	5	32	5	64	768	0.8471	68,706	2,462,850
Model-061	5	8	5	64	768	0.847	106,182	2,423,826
Model-175	5	32	5	32	1024	0.8469	68,706	1,635,170
Model-039	5	8	3	32	1024	0.8468	191,544	1,611,250
Model-125	5	16	5	64	768	0.8462	68,706	2,436,834
Model-187	5	32	5	64	512	0.8457	83,280	1,659,266
Model-114	5	16	3	64	384	0.8453	158,232	1,215,074
Model-047	5	8	5	32	1024	0.8452	77,034	1,615,346
Model-171	5	32	5	32	512	0.8447	104,100	830,818

Co-funded by the European Union

This project has received funding from the European Union’s H2020 research and innovation programme under grant agreement No 680481

Results: Activation Function

	wc1	nc1	wc2	nc2	nfc1	Free Parameters
Model-040	5	8	5	32	128	207730
Model-137	3	32	5	32	256	428130
Model-074	3	16	5	32	384	616258
Model-171	5	32	5	32	512	830818
Model-108	5	16	5	32	640	1018690
Model-114	5	16	3	64	384	1215074
Model-166	5	32	3	32	896	1417698
Model-175	5	32	5	32	1024	1635170
Model-188	5	32	5	64	640	2061058
Model-189	5	32	5	64	768	2462850
Model-062	5	8	5	64	896	2825618
Model-119	5	16	3	64	1024	3224034

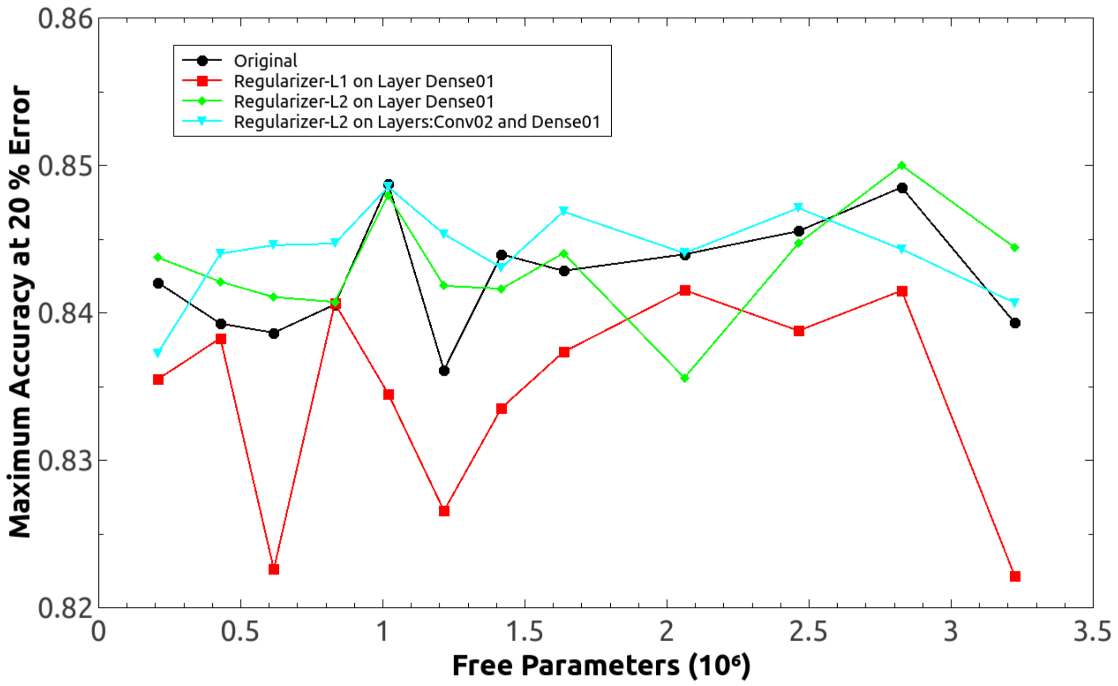


Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

Results: Regularization

	wc1	nc1	wc2	nc2	nfc1	Free Parameters
Model-040	5	8	5	32	128	207730
Model-137	3	32	5	32	256	428130
Model-074	3	16	5	32	384	616258
Model-171	5	32	5	32	512	830818
Model-108	5	16	5	32	640	1018690
Model-114	5	16	3	64	384	1215074
Model-166	5	32	3	32	896	1417698
Model-175	5	32	5	32	1024	1635170
Model-188	5	32	5	64	640	2061058
Model-189	5	32	5	64	768	2462850
Model-062	5	8	5	64	896	2825618
Model-119	5	16	3	64	1024	3224034



$$L1 = \lambda \sum_{i=1}^{k} |w_i|$$

$$L2 = \lambda \sum_{i=1}^{k} w_i^2$$

$$\lambda(L1) = 0.001$$

$$\lambda(L2) = 0.01$$

Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

FORTISSIMO

CESGA

Summary and Conclusions

An hyper-parametric search (# filters, filter size and #neurons) was conducted to obtain best model for monitoring LMD process based on MWIR images.

Thanks to Finis Terrae II HPC CESGA infrastructure the time-to –solution was reduced from 40 days to 70 hours.

Time factor decrease of 14 but could increase to more that 100!!.

Additional approximations like changing activation functions or using regularizers can improve model performance.

Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481

THANKS FOR YOUR ATTENTION !!!

Co-funded by the European Union

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 680481