ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE,TAKES YOU FAR

XXXIV Cycle

# Video Compression using Deep Learning


## Nicola Prette

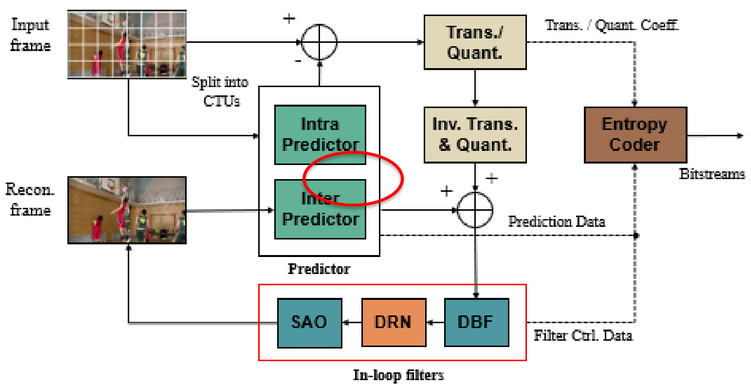
## Supervisor: Profs. Enrico Magli, Tiziano Bianchi


## Research context and motivation

- Developing better and better **Video Compression** algorithms keeps being **a very crucial task**.
- **Deep Learning** techniques have found great success in tackling a multitude of Image and Video processing tasks.
- Reasonable to assume that is possible to design a **Deep Learning** based **Video Compression** algorithm capable of better preserving the perceived visual quality of a video sequence.

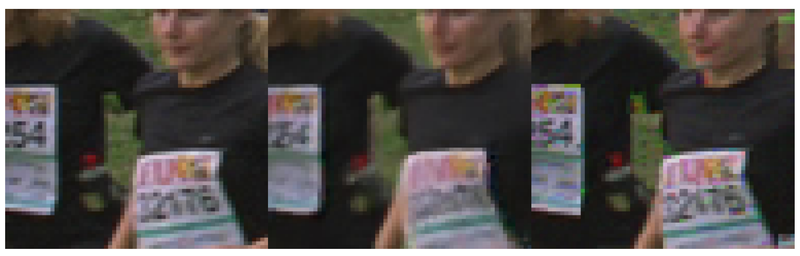
## Addressed research questions/problems

- Difficult to achieve competitive performances compared to currently employed codecs such as **HEVC,** for this reason it was decided to **work inside its framework** and improve upon it using **DL** .
- HEVC employs a **Hybrid Coding Scheme**. Compression is achieved by generating an **estimate of the current frame** and **encoding only the residual**.
- Two kinds of estimations: **Intra prediction,** which is the prediction obtained **exploiting spatial correlations** inside the frame, and **Inter prediction,** which is the prediction obtained exploiting **temporal correlation** between frames.
- Inter-Prediction is done by estimating the **spatial displacements (Motion Vectors)** between patterns in the figure**, which are then used to** obtain the estimated frame.
- The downside is that these **Motion Vectors** have to be encoded together with the residuals to allow the correct reconstruction of the compressed video.
- The aim of the research is to create a **Deep Learning** algorithm which is capable of generating an estimate of **future frames** without needing **Motion Vectors** and thus reducing the amount of information needed for the **video reconstruction**.




## Partial Results

- The **Network** was trained on **60000**, **4 frames** sequences of dimension **64x64 pixels** and was made to generate a frame given **3 previous ones**.
- Same performanes in terms of **PSNR** of classical motion compensation techniques .



Ground Truth | Filter Generating Network | Motion Compensation

## Submitted and published works

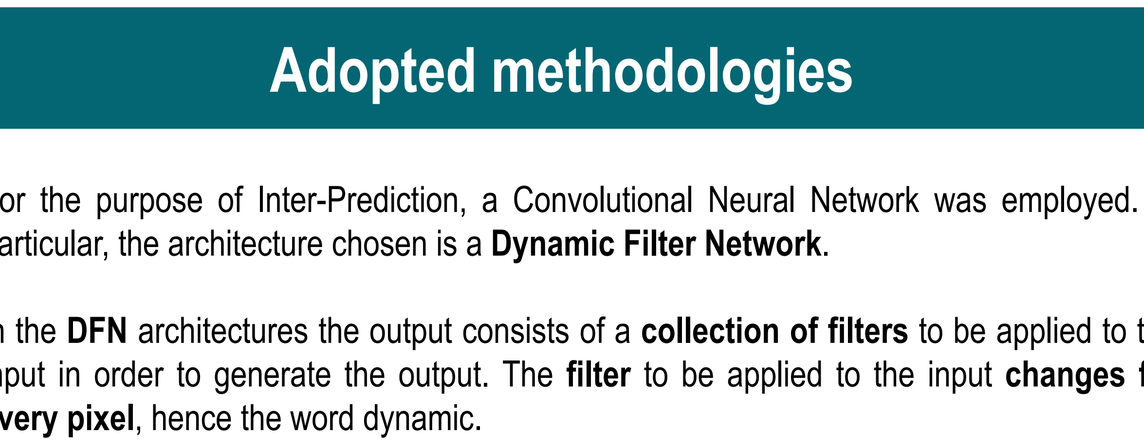

- Prette, Nicola; Magli, Enrico; and Bianchi, Tiziano., "*Using CCSDS image compression standard for SAR raw data compression in the H2020 EO-ALERT project*", 2019 European Workshop on On-board Data processing, 2019


## Novel contributions

- This work aims at creating a viable way to integrate machine learning inside the **current Video compression framework** in order to improve the performances in terms of **Rate Distortion** without increasing the complexity of the algorithm.
- The focus was placed on improving the **Intra Prediction** process as to not require the transmission of motion vector **without increasing the magnitude of the residuals**.
- The adopted architecure learns to model motion with a straight forward convolutional network structure, no recursion or memory needed

## Adopted methodologies

- For the purpose of Inter-Prediction, a Convolutional Neural Network was employed. In particular, the architecture chosen is a **Dynamic Filter Network**.
- In the **DFN** architectures the output consists of a **collection of filters** to be applied to the input in order to generate the output. The **filter** to be applied to the input **changes for every pixel**, hence the word dynamic.
- This kind of **architecture is well suited** for tackling **frame prediction** as it was previously employed for the purpose of **video super resolution** and it has been shown to be able to **learn motion models** without explicitly computing **motion vectors**.




## Future work

- Further improvement of the quality of the **Inter-Prediction** and reduction of the **blurriness** of the predicted frames.
- **Implementation** of the developed **Inter-Prediction** algorithm inside the architecture of **HEVC** and analisys of the performances in terms of **Rate/Distortion** after the modification.
- Extention of the **Deep Learning** approach to other phases of the **compression algorithm**.

## List of attended classes

- 02LWHRV – Communication (22/12/2018, 1 credit)
- 01RNBRV – Communication II (03/05/2019, 2 credits)
- 01QRQRV – Compressed sensing: theory and applications (30/05/2019, 4 credits)
- 01TEHRV – Data Science for Networks (15/02/2019, 6 credits)
- 01QTEIU – Data mining concepts and algorithms (14/12/2018, 4 credits)
- 01SFURV – Programmazione scientifica avanzata in matlab (15/05/2019, 4 credits)
- 01RNCRV – Public Speaking II (03/05/2019, 2 credits)
- 01RISRV – Public speaking (27/11/2018, 1 credits)
- 01SYBRV – Research integrity (13/12/2018, 1 credits)
- 01SWPRV – Time management (17/12/2018, 1 credits)
- 01TBXRV – Vision fundamentals in service robotics (23/05/2019, 4 credits)
- 01QORRV – Writing Scientific Papers in English (28/03/2019, 3 credits)

POLITECNICO DI TORINO

PhD program in
**Electrical, Electronics and Communications Engineering**