

SUMMARIZATION OF DEEP LEARNING PAPER

**JENNIFER CACERES
COMPUTER SCIENCE - MASTER PROGRAM
CITY OF NEW YORK**



REAL-TIME SINGLE IMAGE AND VIDEO SUPER-RESOLUTION USING AN EFFICIENT SUB-PIXEL CNN

BY:

**HARRISON JESSE SMITH
QINGYUAN ZHENG**

YIFEI LI

SOMYA JAIN, & JESSICA K. HODGINS

PROBLEM

- single-image and video super-resolution
- Includes:
 - enhancing the quality of low-res images & videos
 - reconstructing missing details
 - generating higher-res versions



IMPORTANCE

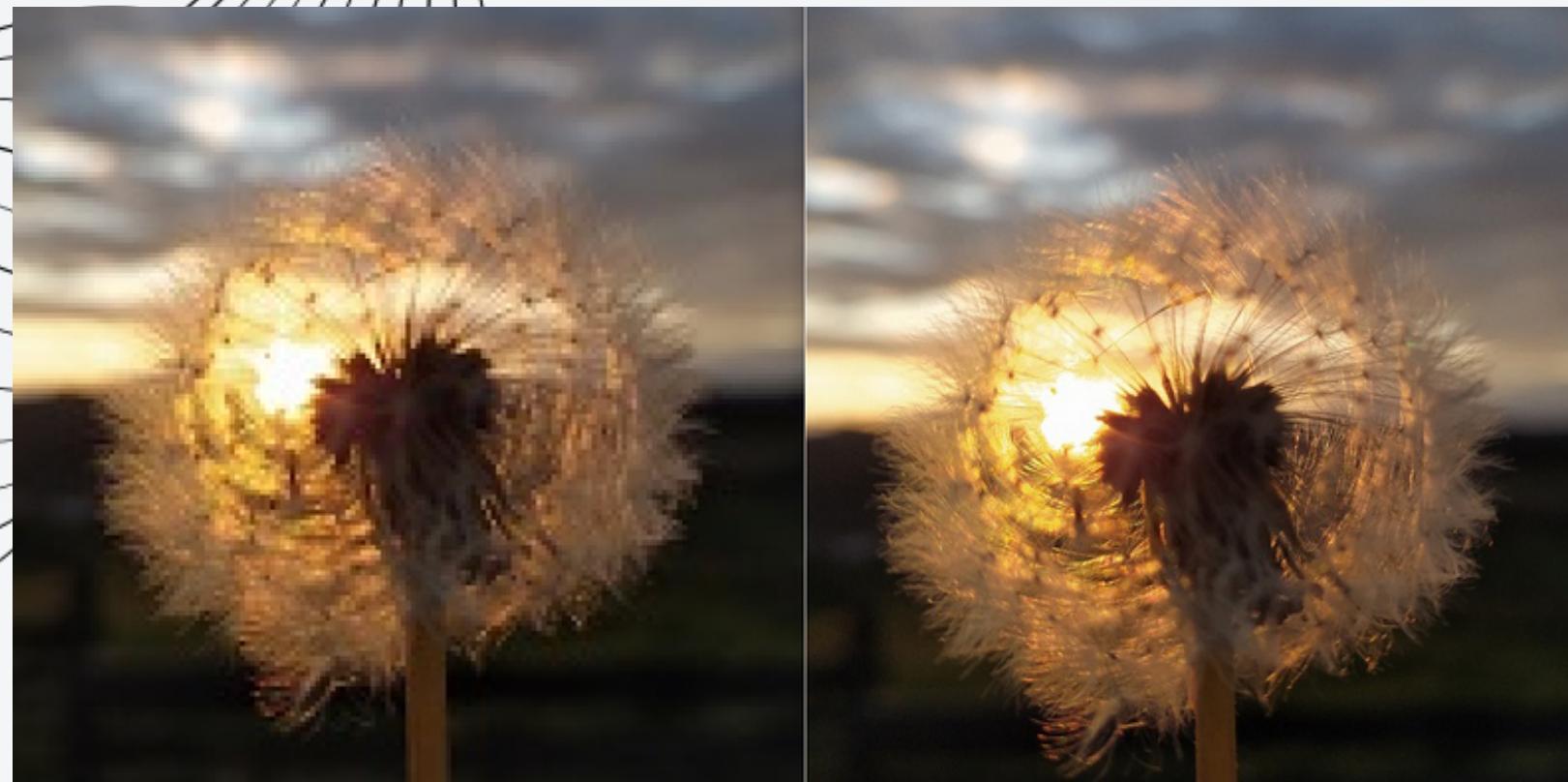
- Super-resolution is a curial problem in various domains, like:
 - surveillance
 - medical imaging
 - entertainment



IMPACT

- Solving this problem can lead to;
 - Better Visualization and analysis of low-res images & videos
 - improving the performance of computer vision algorithms
 - improving Real-time super-resolution techniques





High Resolution



Low Resolution



**WHAT WAS
ACCOMPLISH?**

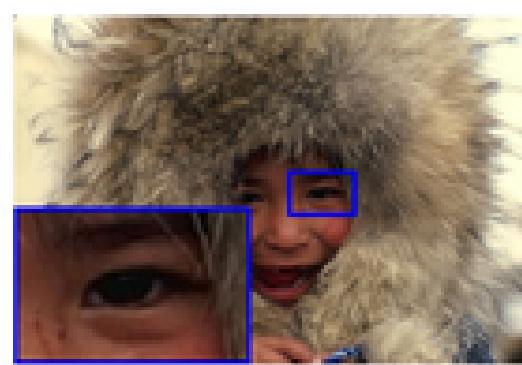
ACCOMPLISHMENTS

Here are some of the accomplishments that were made by the authors:

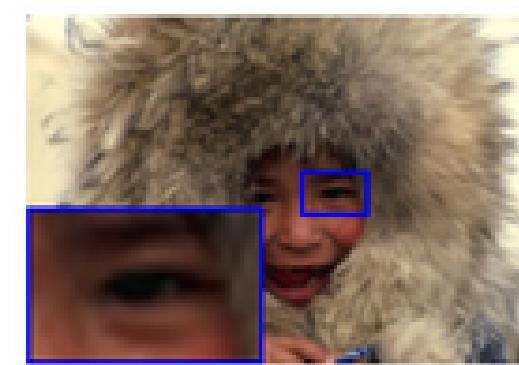
- Real-time Super-Resolution
- Improved Visual Quality
- Easy Implementation

REAL-TIME SUPER-RESOLUTION

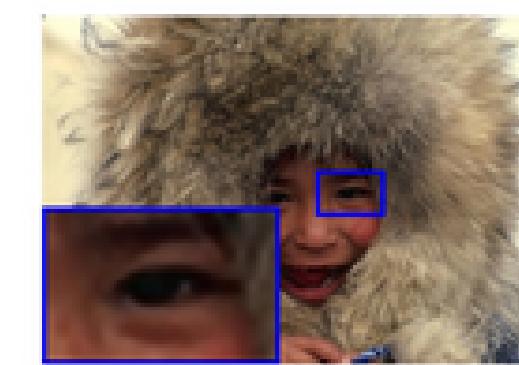
Using a sub-pixel convolutional
neural network (CNN) that
performs a single-image.



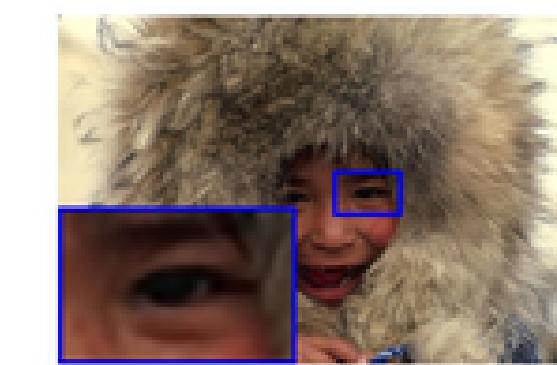
(a) 14092 Original



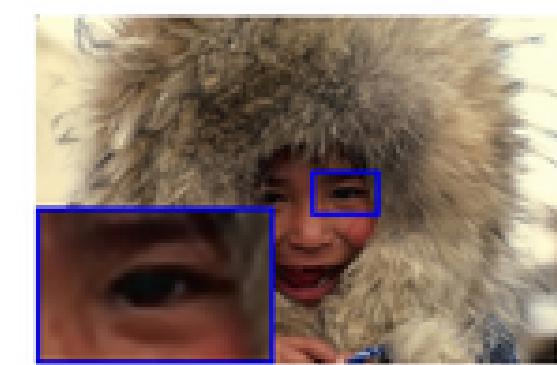
(b) Bicubic / 29.06db



(c) SRCNN [7] / 29.74db



(d) TNRD [3] / 29.74db



(e) ESPCN / **29.78db**

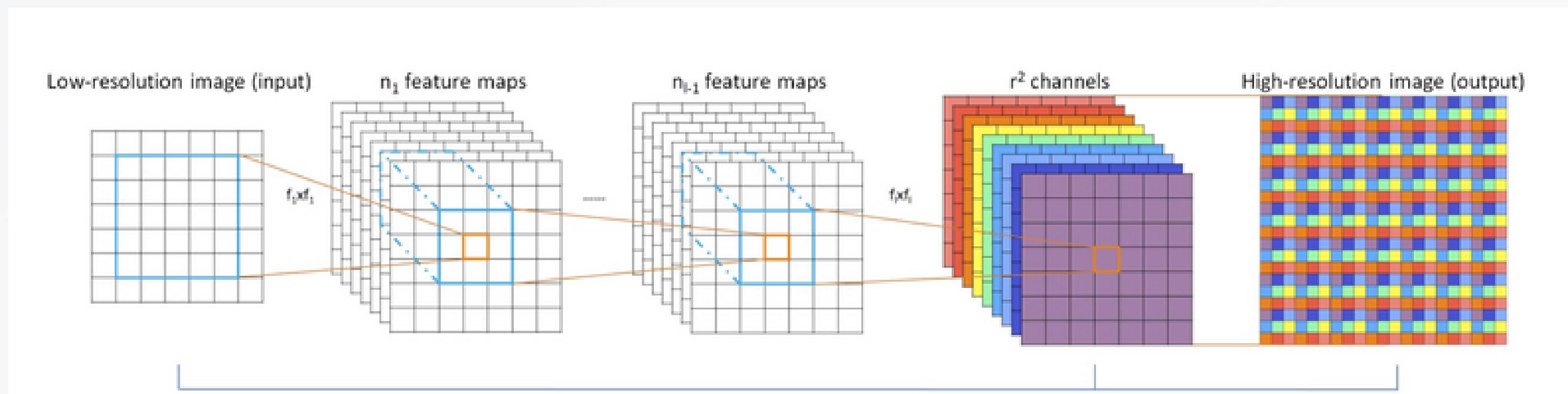
IMPROVED VISUAL QUALITY

Enhances the visual quality of low-resolution images by reconstructing missing details and generating sharper

Dataset	Scale	SRCNN (91)	ESPCN (91 <i>relu</i>)	ESPCN (91)	SRCNN (ImageNet)	ESPCN (ImageNet <i>relu</i>)
Set5	3	32.39	32.39	32.55	32.52	33.00
Set14	3	29.00	28.97	29.08	29.14	29.42
BSD300	3	28.21	28.20	28.26	28.29	28.52
BSD500	3	28.28	28.27	28.34	28.37	28.62
SuperTexture	3	26.37	26.38	26.42	26.41	26.69
Average	3	27.76	27.76	27.82	27.83	28.09

EASY IMPLEMENTATION

provide code and implementation details, enabling easy replication and utilization of their super-resolution technique.





**THERE'S
MORE ...**

PREVIOUS WORK

There are various techniques proposed over the years, like

- Traditional interpolation methods
- Sparse coding-based approaches
- other deep learning based methods

The ESPCN model was introduced as a novel deep learning based solution.

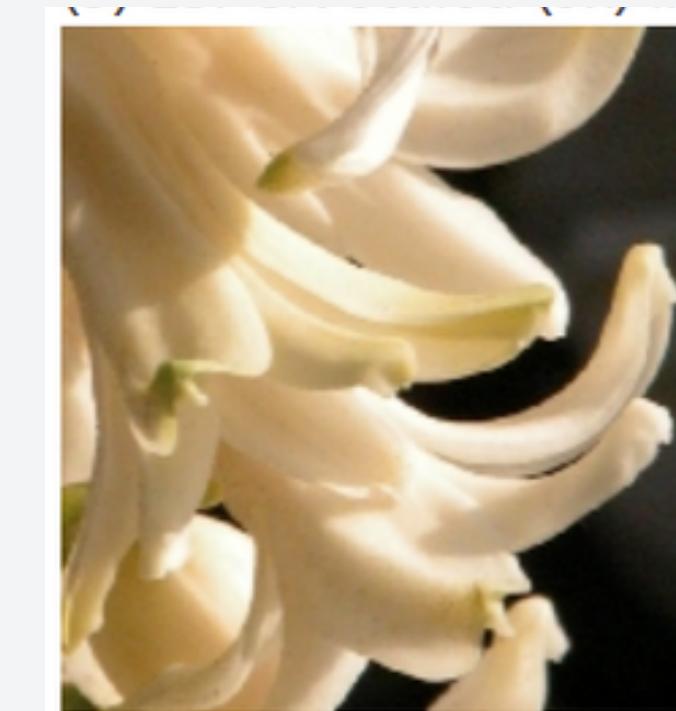
- Convolutional neural networks (CNNs)
- Real-time processing



**HOW DOES IT FIT
TOGETHER?**

INPUT PREPROCESSING

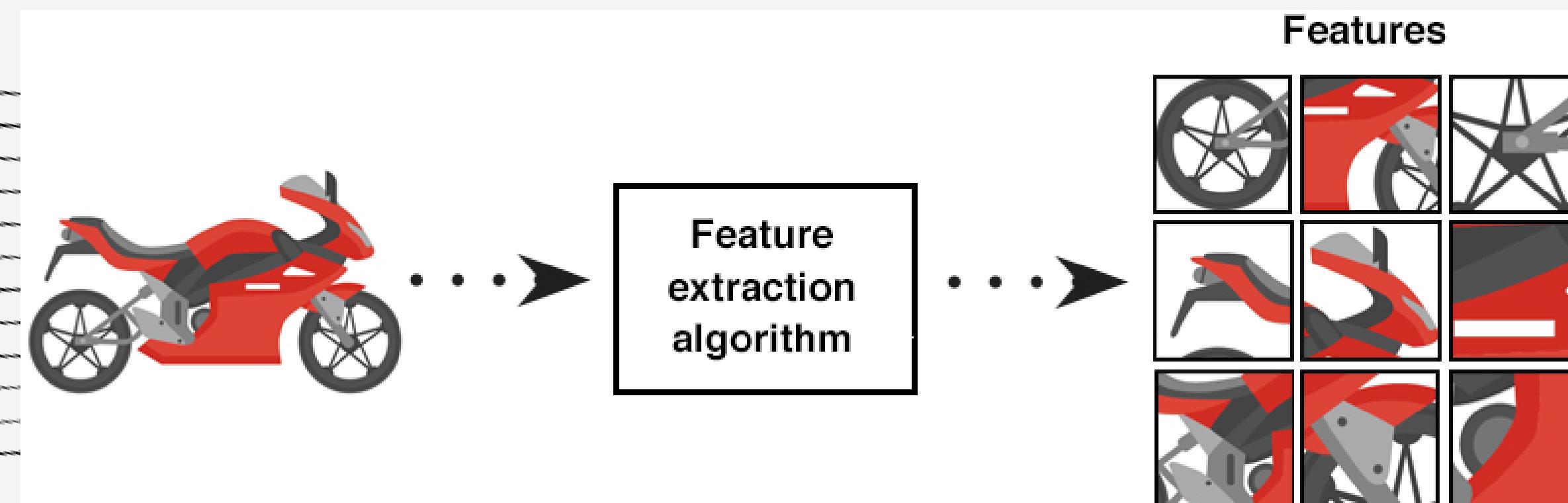
The model takes low-resolution images as input. These low-resolution inputs are typically down sampled as version of the original high-resolution content



FEATURE EXTRACTION

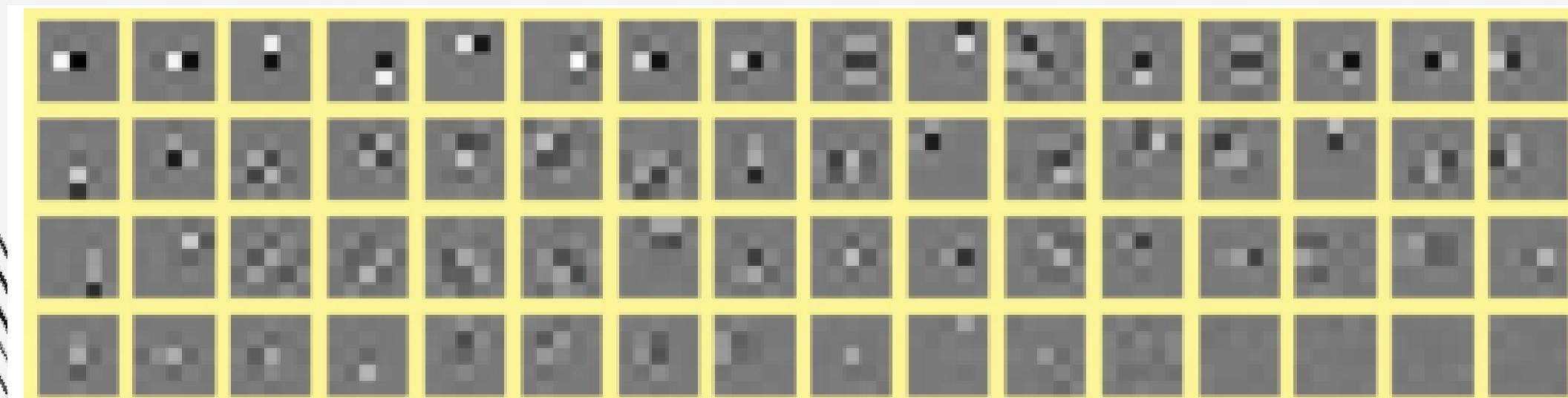
The ESPCN network with the series of CNN layers that extract meaningful features from the post-process input.

The goal is to capture relevant patterns and details that can contribute to the enhancement of the output.



NON-LINEAR MAPPING

Using non-linear activation function, such as ReLU into the model and enable it to learn complex relationships between low vs high.



LEARNING

Has sub-pixel convolutional layers that perform
upsampling within the network itself

RECONSTRUCTING

The final layers of the network focus on reconstructing
the enhanced high-resolution image

LOSS FUNCTION AND TRAINING

Loss Function:

the difference between the generated high-resolution output and ground truth high-resolution data

RESULTS

- (1) Original picture
- (2) Bicubic scaled (3x) image
- (3) ESPCN scaled (3x) image



**THANK'S FOR
WATCHING**

