Higher-order Statistical Modeling based Deep CNNs (Introduction)

Peihua Li

Dalian University of Technology

http://peihuali.org

Outline

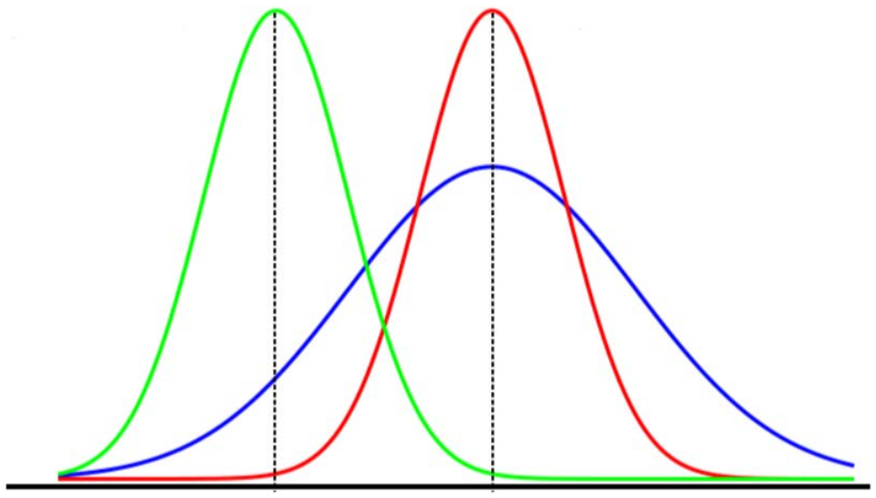
- What is Higher-order?
- Why We Study High-order
- Overview of Speaker
- Overview of Tutorial

What is Higher-order?—Statistical Moments

For **scalar** random variable X, $f_X(x)$ is its proability density

1st-order moment $$E(X) = \int_{\mathbb{R}} x f_X(x) dx$$

2nd-order moment $$E(X^2) = \int_{\mathbb{R}} x^2 f_X(x) dx$$



kth-order moment $$E(X^k) = \int_{\mathbb{R}} x^k f_X(x) dx \quad E(X^k) = \frac{1}{N}\sum_i x_i^k$$

What is Higher-order?—Statistical Moments

For **scalar** random variable X, $f_X(x)$ is its proability density

i.i.d. samples

1st-order moment $E(X) = \int_{\mathbb{R}} x f_X(x) dx$ $E(X) = \frac{1}{N} \sum_i x_i$

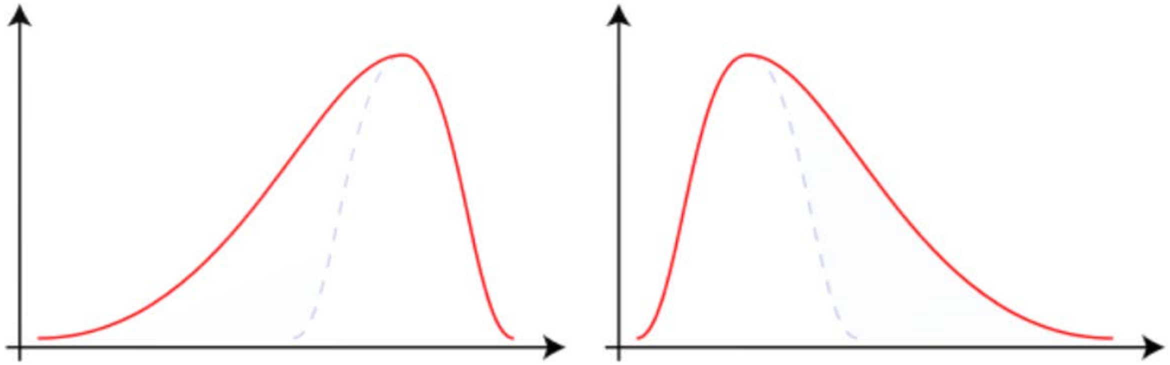
2nd-order moment $E(X^2) = \int_{\mathbb{R}} x^2 f_X(x) dx$ $E(X^2) = \frac{1}{N} \sum_i x_i^2$

kth-order moment $E(X^k) = \int_{\mathbb{R}} x^k f_X(x) dx$ $E(X^k) = \frac{1}{N} \sum_i x_i^k$

What is Higher-order?—Statistical Moments

For **scalar** random variable X, $f_X(x)$ is its proability density

i.i.d. samples



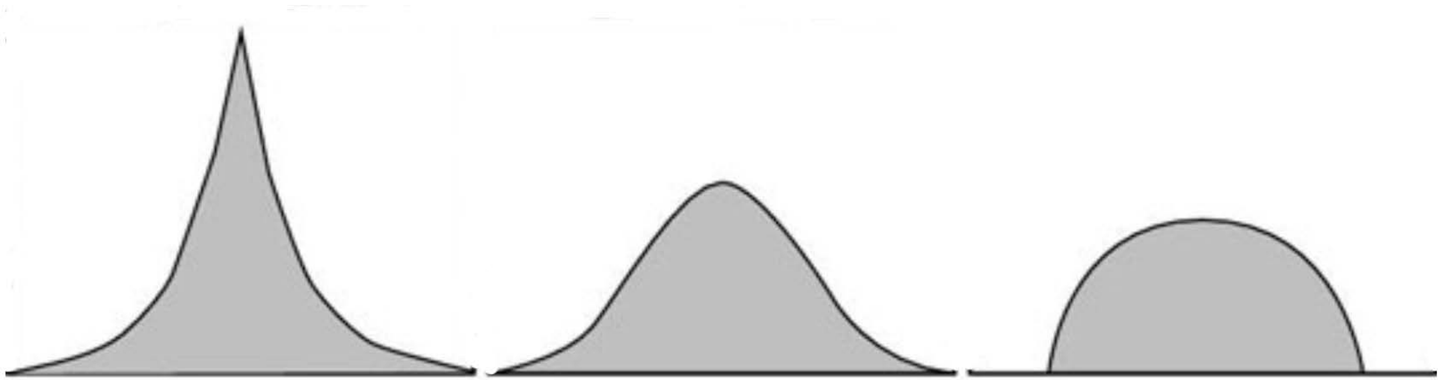
3rd-order moment

k^{th}-order moment $$E(X^k)=\int_{\mathbb{R}} x^k f_X(x)dx \quad E(X^k)=\frac{1}{N}\sum_i x_i^k$$

What is Higher-order?—Statistical Moments

For **scalar** random variable X, $f_X(x)$ is its proability density

i.i.d. samples

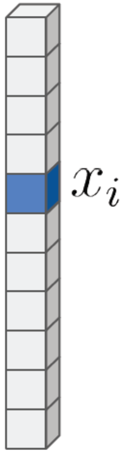


4th-order moment

*k*th-order moment $E(X^k)=\int_{\mathbb{R}} x^k f_X(x)dx$ $E(X^k)=\frac{1}{N}\sum_i x_i^k$

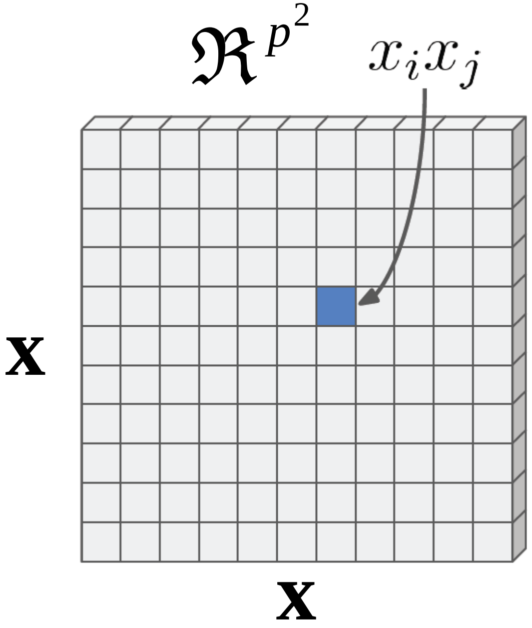
What is Higher-order?—Statistical Moments

Random vector $\mathbf{x} \in \mathfrak{R}^{p}$



1st-order moment

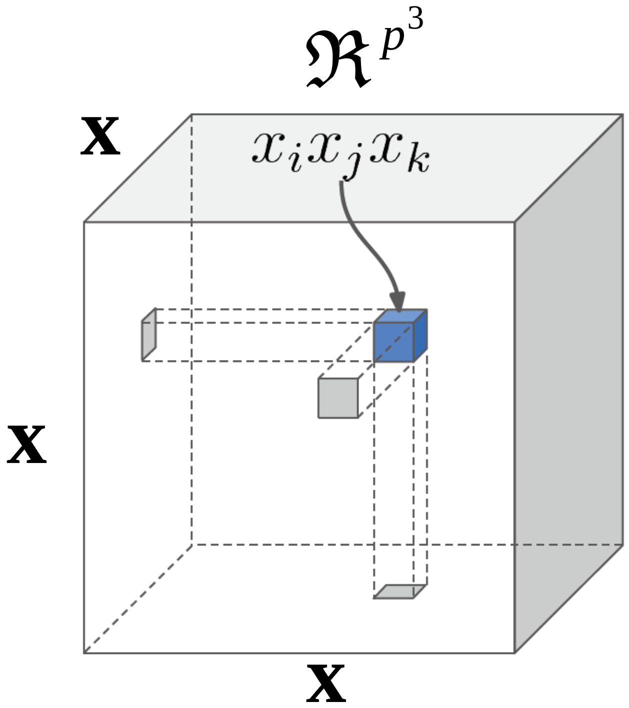
$$E(X) = \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{C}} \mathbf{x}$$



2nd-order moment

$$E(X^2) = \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{C}} \mathbf{x}\mathbf{x}^T$$

$$= \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{C}} \mathbf{x} \otimes \mathbf{x}$$



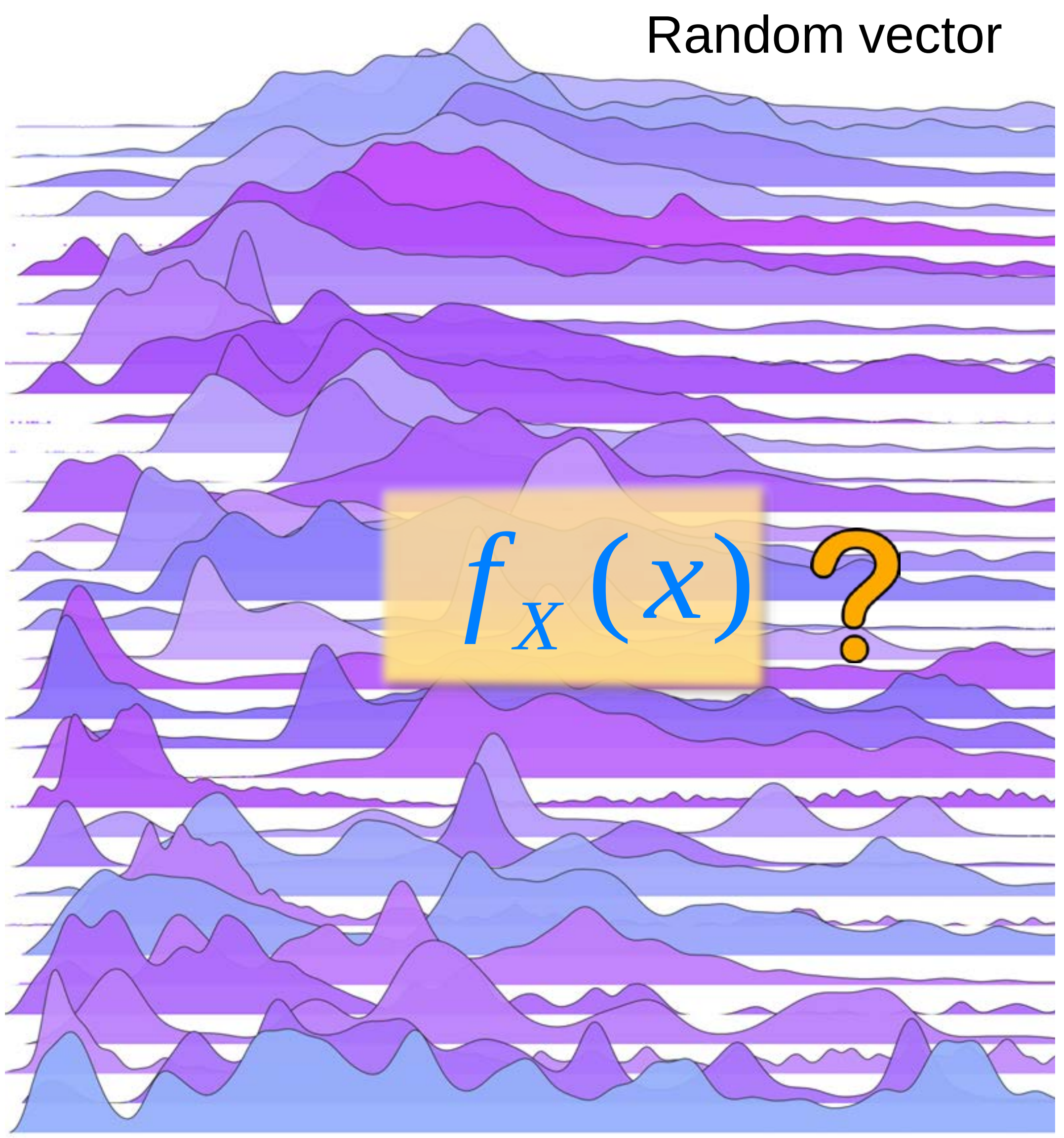
3rd-order moment

$$E(X^3) = \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{C}} \mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}$$

Images courtesy of "Kernel Pooling for Convolutional Neural Networks"

What is Higher-order?—Statistical Moments

Probaiblity density $f_X(x)$ is everything



What is Higher-order?—Statistical Moments

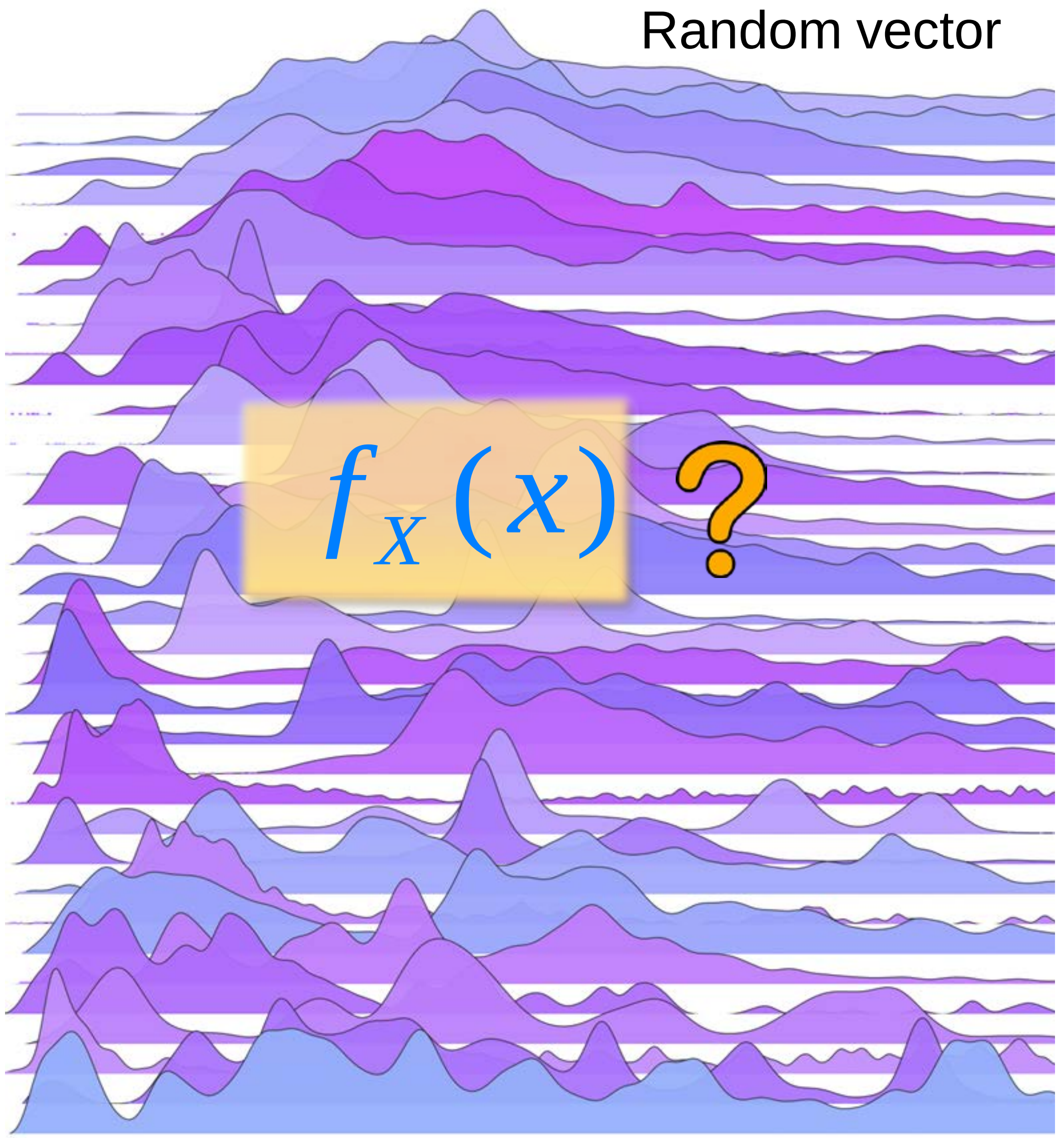
Characteristic Function=Probability Density

The characteristic function is defined as

$$\Phi_X(\omega) = E\left(e^{jX\omega}\right) = \int_{-\infty}^{+\infty} f_X(x) e^{jx\omega} dx$$

Fouriere Transform Pair

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi_X(\omega) e^{-jx\omega} dx$$



What is Higher-order?—Statistical Moments

If we know characteristic function $\Phi_X(\omega)$, we know everything.

$$f_X(x)$$

Fouriere Transform Pair

$$\Phi_X(\omega) = E\left(e^{jX\omega}\right)$$

Moments matter

$$= E\left[\sum_{k=0}^{\infty} \frac{(j\omega X)^k}{k!}\right] = \sum_{k=0}^{\infty} \frac{j^k}{k!} E(X^k)\omega^k$$

$$= 1 + jE(X)\omega + \frac{j^2}{2!}E(X^2)\omega^2 + \cdots + \frac{j^k}{k!}E(X^k)\omega^k + \cdots.$$

What is Higher-order?—Statistical Moments

If we know characteristic function $\Phi_X(\omega)$, we know everything.

Moments matter

$$f_X(x) \Longleftrightarrow \Phi_X(\omega) = \sum_{k=0}^{\infty} \frac{j^k}{k!} E(X^k) \omega^k$$

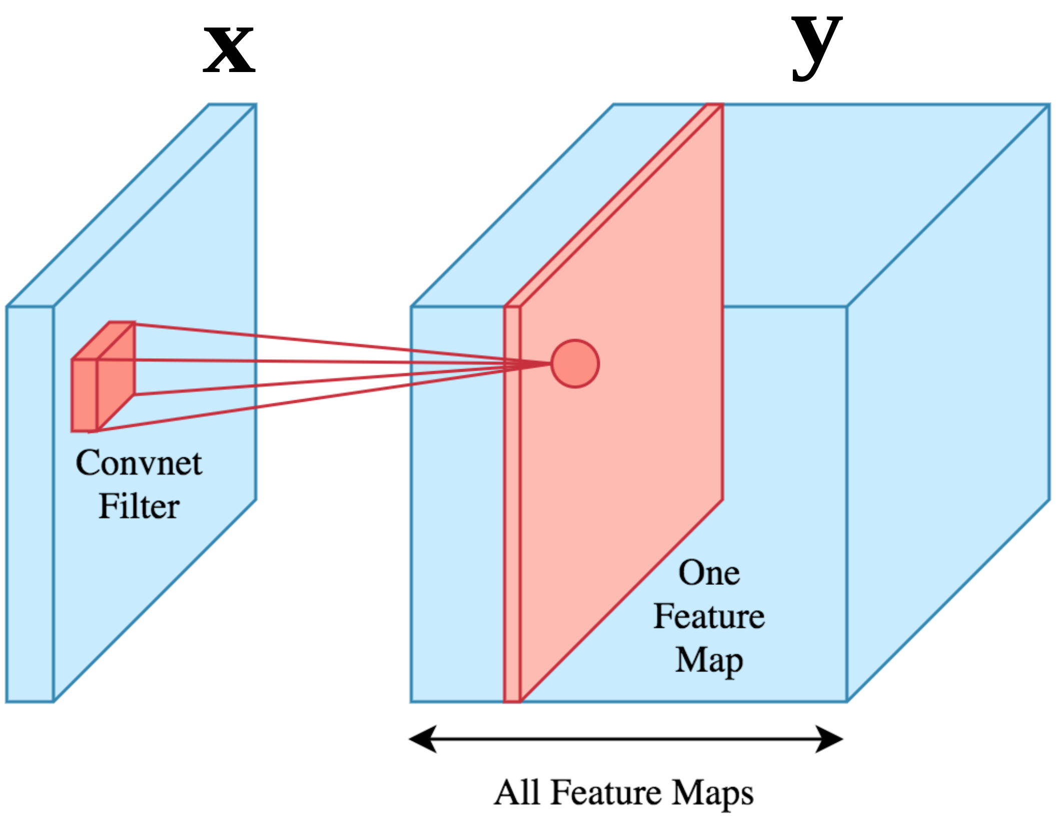
Probability density **Fouriere Transform** **Characteristic function**

1st-order moment $$E(X) = \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{C}} \mathbf{x}$$

2nd-order moment $$E(X^2) = \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{C}} \mathbf{x}\mathbf{x}^T = \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{C}} \mathbf{x} \otimes \mathbf{x}$$

3rd-order moment $$E(X^3) = \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{C}} \mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}$$

What is Higher-order? — Signal Perspective



$$\mathbf{y} = f(\mathbf{x}) \;?$$

Convolution is a linear transformation

$$\mathbf{y} = \mathbf{b} + \mathbf{W}\mathbf{x}$$

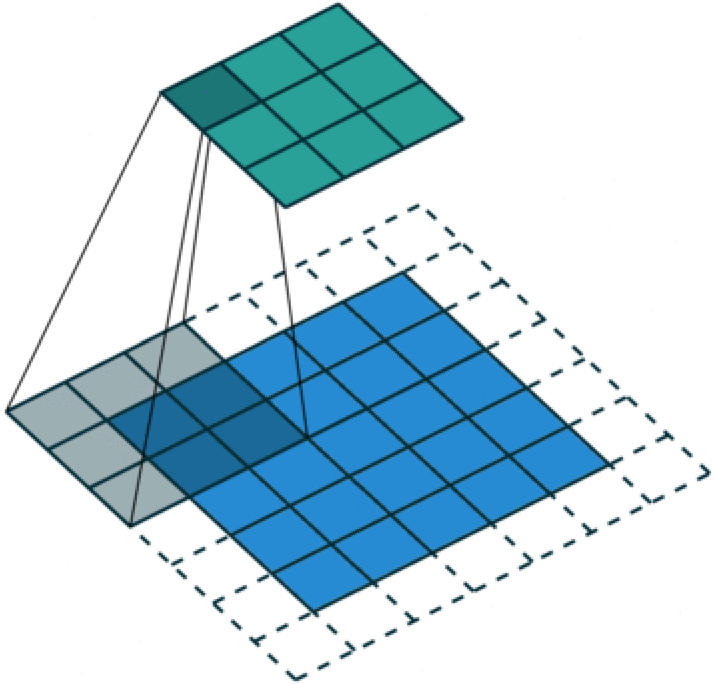
Multi-variable Taylor series:

$$\mathbf{y} = \mathbf{b} + \mathbf{W}\mathbf{x} + \mathbf{H}(\mathbf{x} \otimes \mathbf{x}) + \cdots$$

1st-order term
(Linear term)

2nd-order term
(Quadratic term)

What is Higher-order?—Signal Perspective



Convolution is a linear transformation

Two variable Taylor series:

$$f(u,v) = f(0,0) + \mathbf{w}^T \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} u & v \end{bmatrix} \mathbf{H} \begin{bmatrix} u \\ v \end{bmatrix} + \cdots$$

Linear term

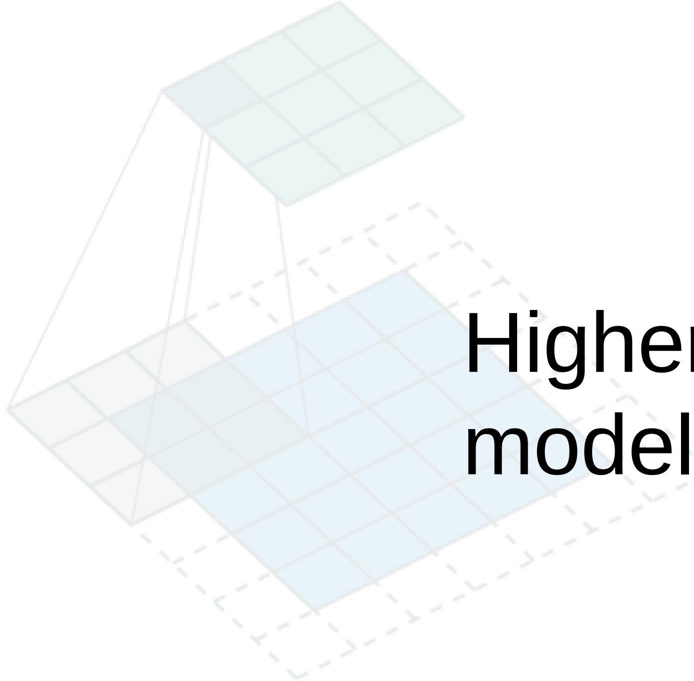
$$w_1 u + w_2 v$$

Quadratic term

$$h_{11}u^2 + 2h_{12}uv + h_{22}v^2$$

What is Higher-order?—Signal Perspective

Higher order enhances non-linear modeling capability



Convolution is a linear transformation

Two variable Taylor series:

$$f(u,v) = f(0,0) + \mathbf{w}^T \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} u & v \end{bmatrix} \mathbf{H} \begin{bmatrix} u \\ v \end{bmatrix} + \cdots$$

Linear term

$$w_1 u + w_2 v$$

Quadratic term

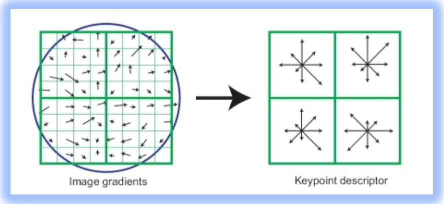
$$h_{11}u^2 + 2h_{12}uv + h_{22}v^2$$

Outline

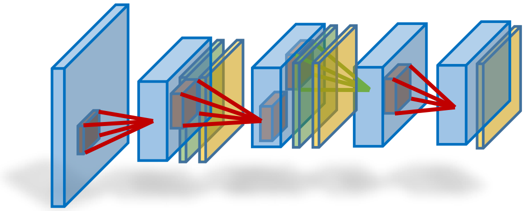
- What is Higher-order?
- Why We Study High-order
- Overview of Speaker
- Overview of Tutorial

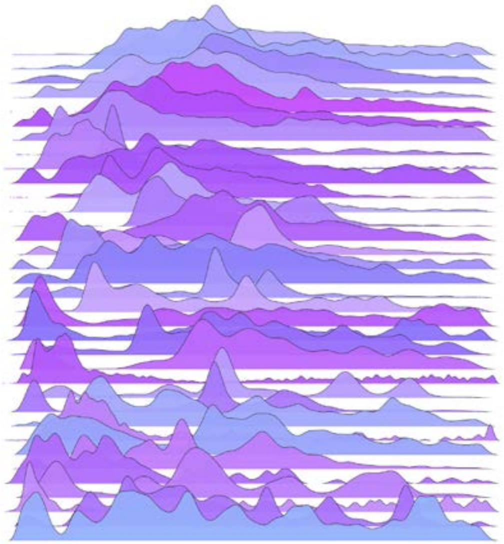
Why Higher-order?

Hand-crafted Features



Learned Features





$$f_X(x) \Leftrightarrow \Phi_X(\omega) = \sum_{k=0}^{\infty} \frac{j^k}{k!} E(X^k)\omega^k$$

Probability density

Characteristic function

Why Higher-order?

Hand-crafted Features

Higher-order moments can better characterize real-word distributions

Learned Features

$$f_X(x) \Leftrightarrow \Phi_X(\omega) = \sum_{k=0}^{\infty} \frac{j^k}{k!} E(X^k)\omega^k$$

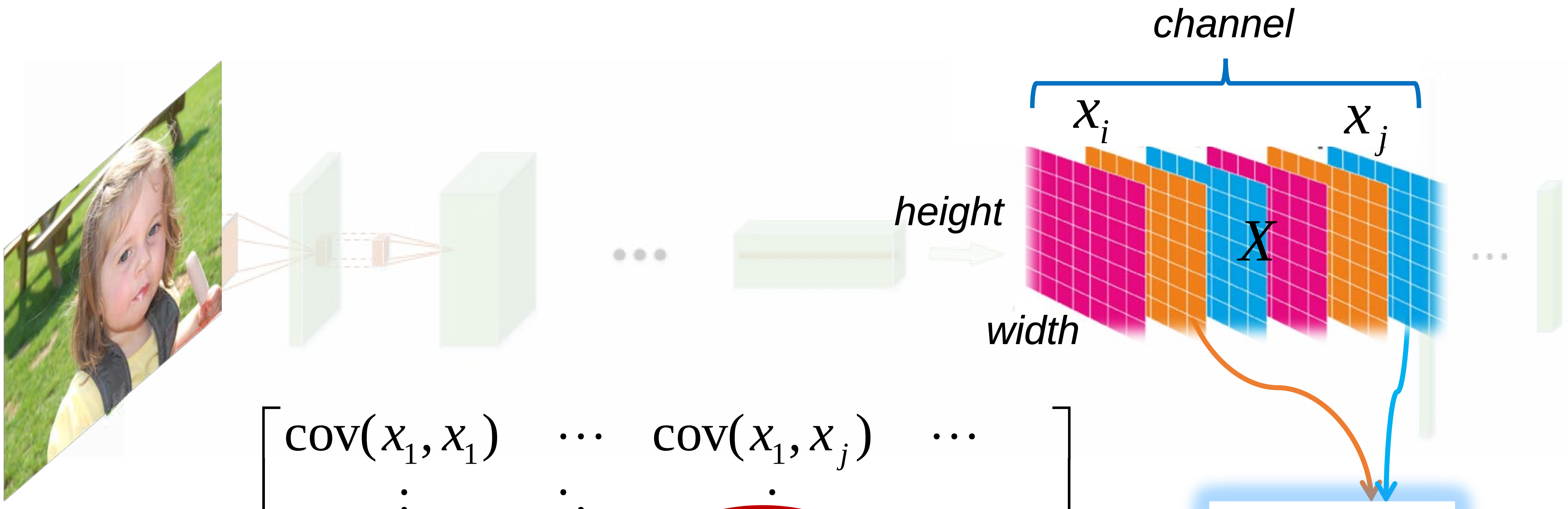
Probability density

Characteristic function

Why Higher-order?

2nd-order moment

$$E(X^2) = \frac{1}{N}\sum_{\mathbf{x}\in\mathbb{C}} \mathbf{x}\mathbf{x}^T$$

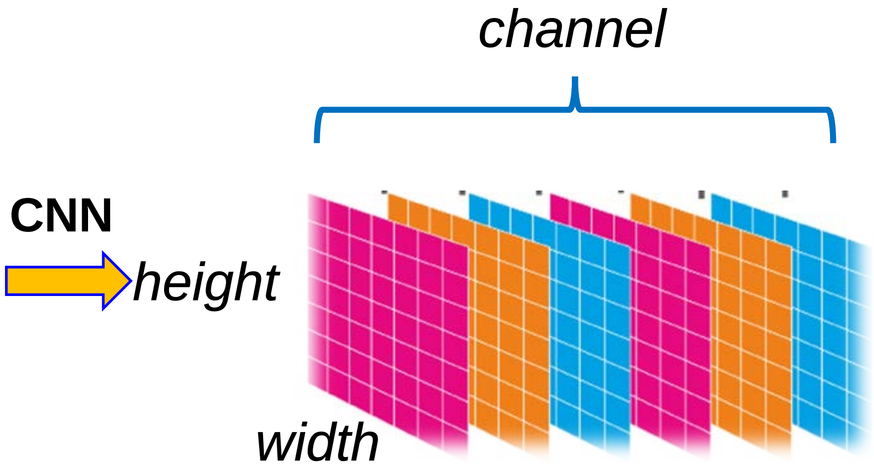


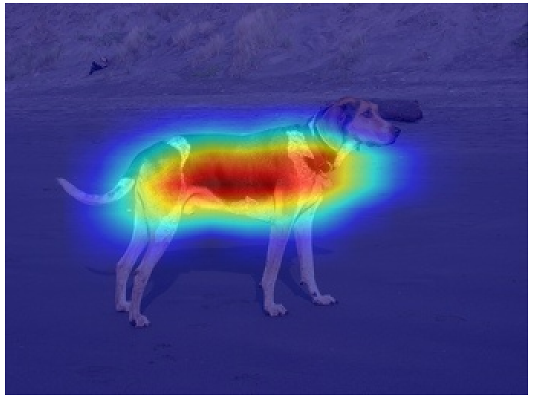
$$\begin{bmatrix} \mathrm{cov}(x_1, x_1) & \cdots & \mathrm{cov}(x_1, x_j) & \cdots \\ \vdots & \ddots & \vdots & \\ \mathrm{cov}(x_i, x_1) & \cdots & \mathrm{cov}(x_i, x_j) & \cdots \\ \vdots & & \vdots & \\ \mathrm{cov}(x_p, x_1) & \cdots & \mathrm{cov}(x_p, x_j) & \cdots \end{bmatrix}$$

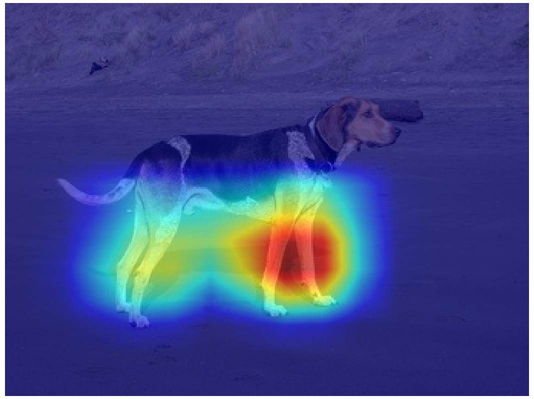
channel correlation

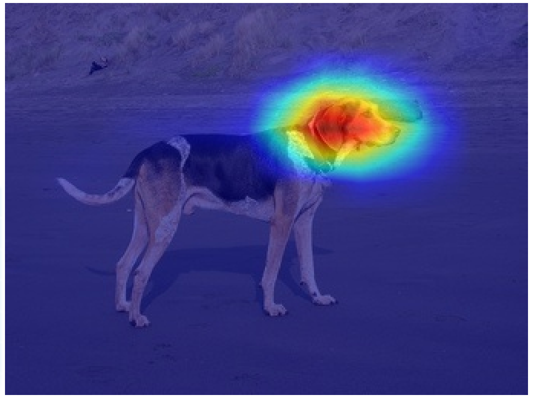
Why Higher-order?

What does each channel indicate?

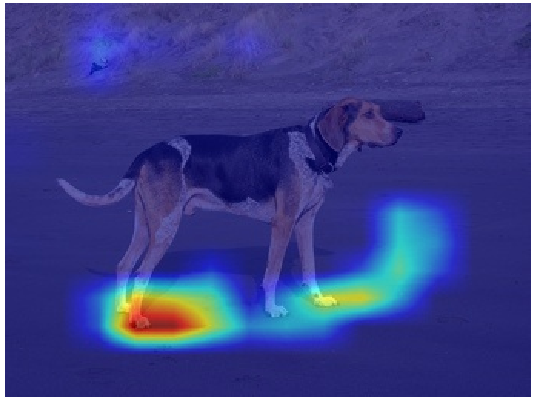



Body | channel 452


Legs | channel 99


Head | channel 123


Tail | channel 174


Hind claw | channel 448


Front claw | channel 333

B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. Learning Deep Features for Discriminative Localization. Computer Vision and Pattern Recognition (CVPR), 2016.

Why Higher-orc

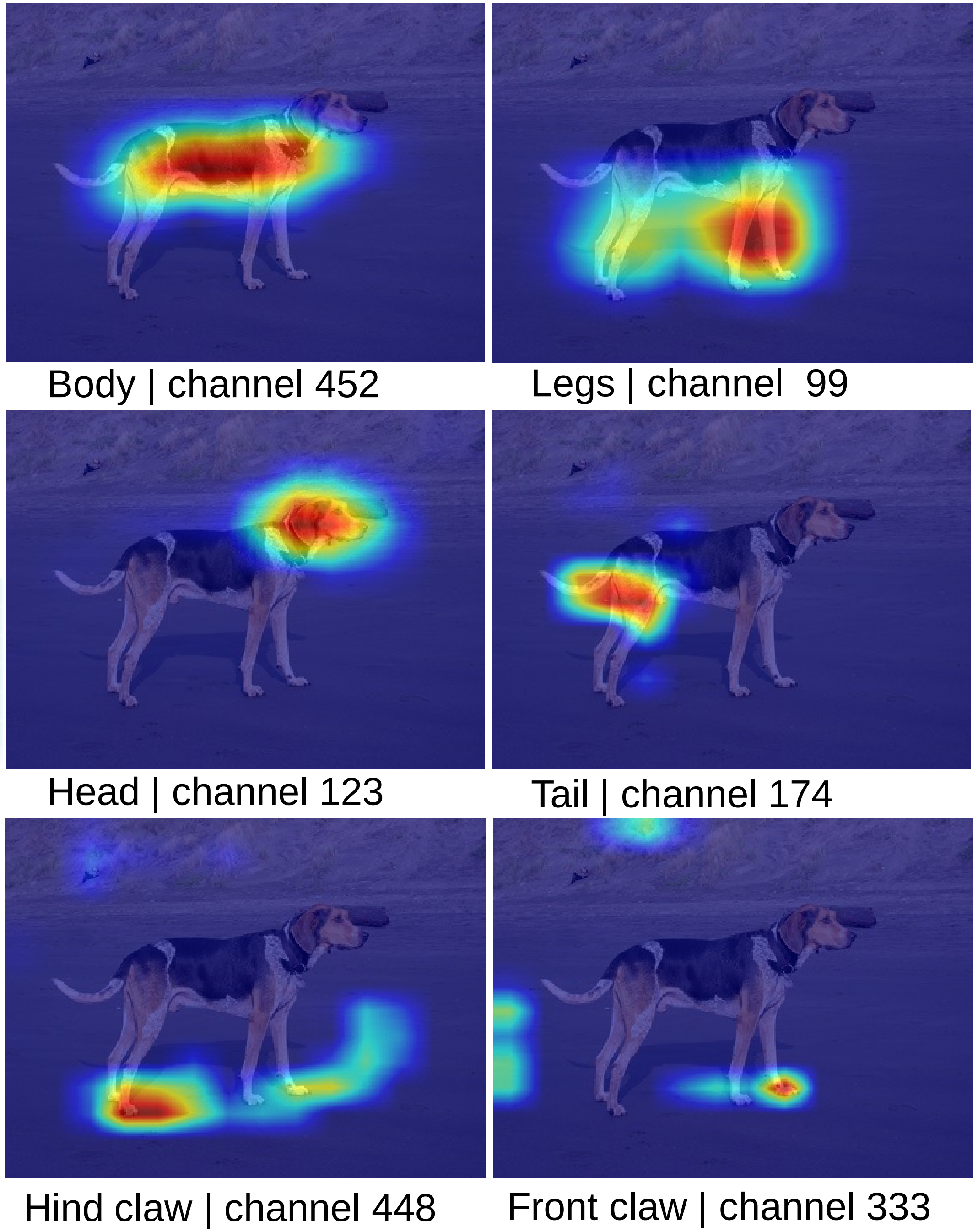
What does each channe

Physical Interpretation

For object recogniton, 2nd-order moment capture dependency of different parts

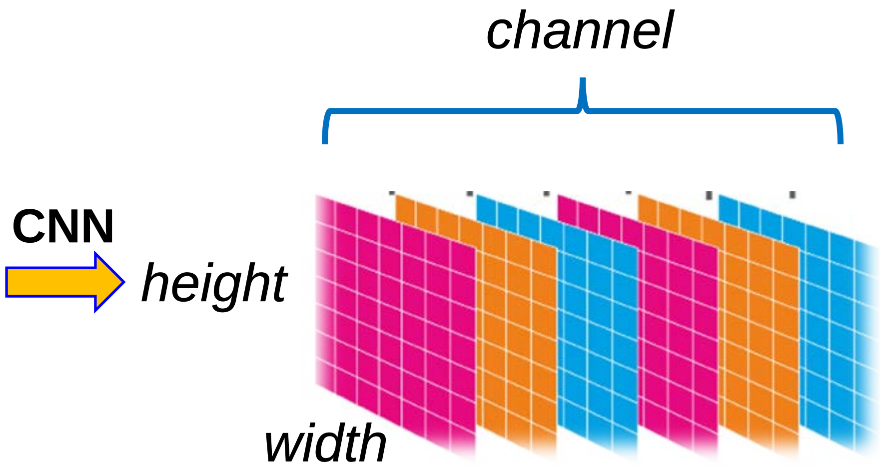
Context of the object

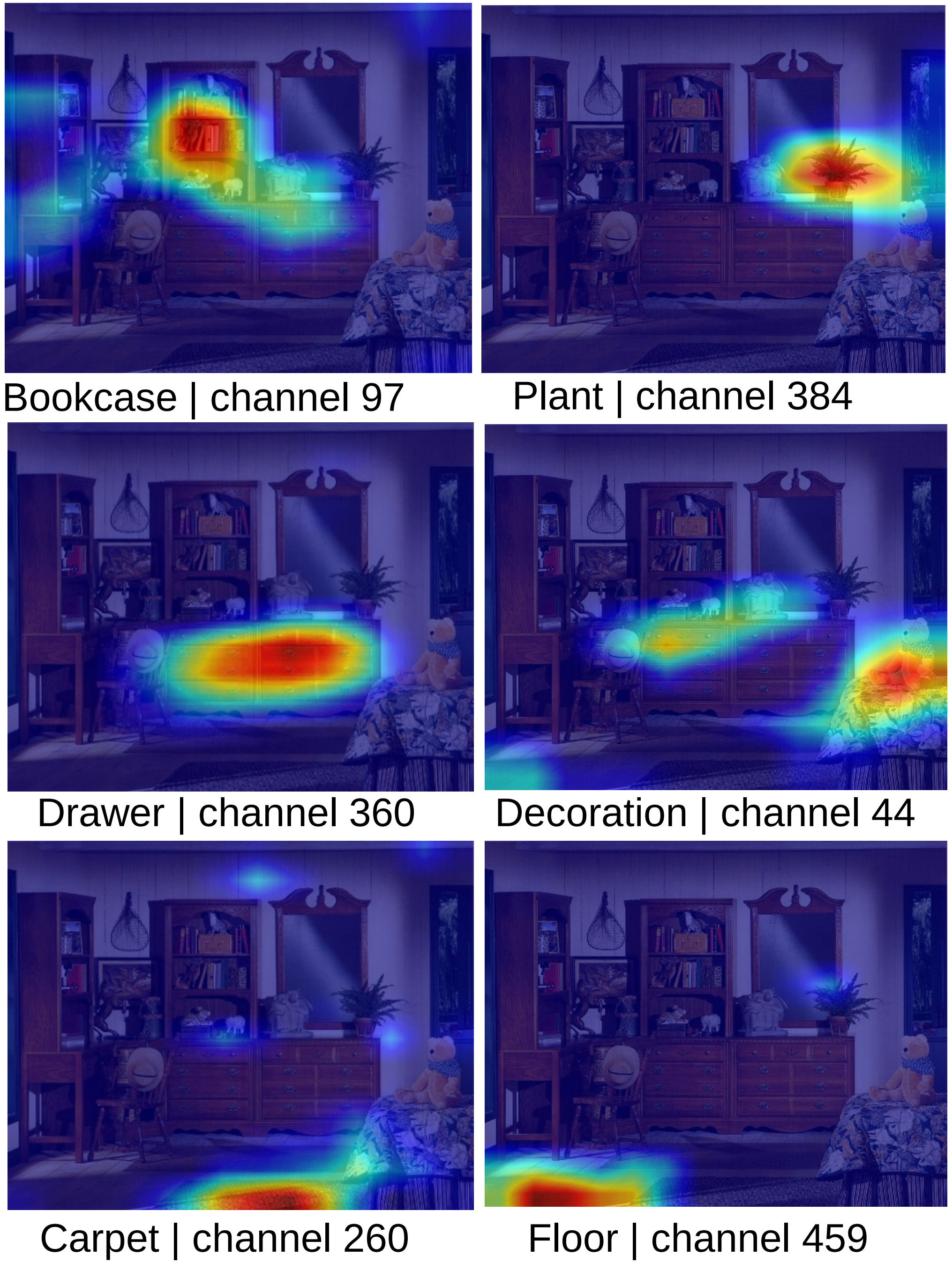
B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. Learning Deep Features for Discriminative Localization. Computer Vision and Pattern Recognition (CVPR), 2016.



Why Higher-order?

What does each channel indicate?





B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. Learning Deep Features for Discriminative Localization. Computer Vision and Pattern Recognition (CVPR), 2016.

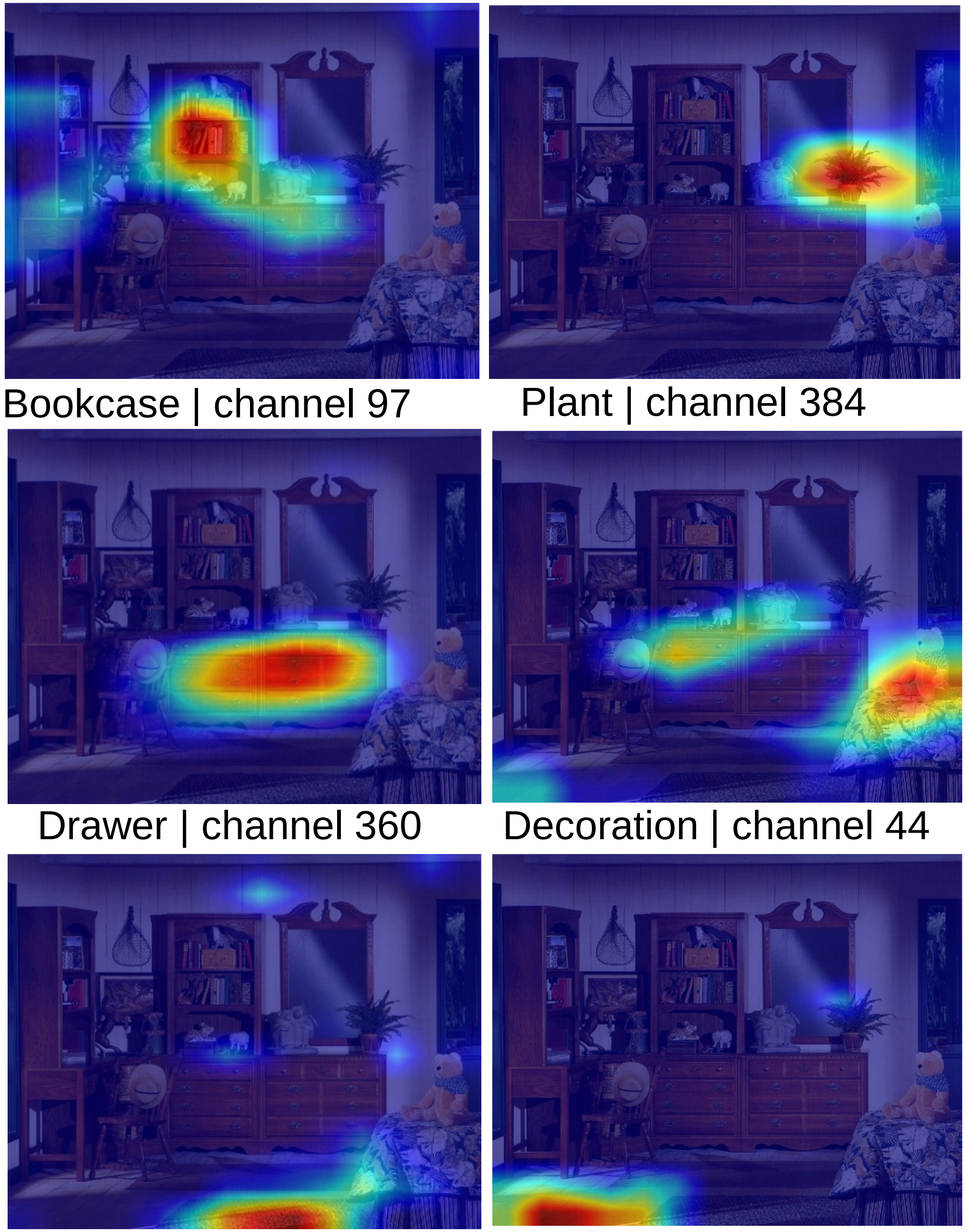
Why Higher-ord

What does each chann

Physical Interpretation

For scene images, 2^{nd}-order moment capture dependency of different objects

Context of the scene



Bookcase | channel 97

Plant | channel 384

Drawer | channel 360

Decoration | channel 44

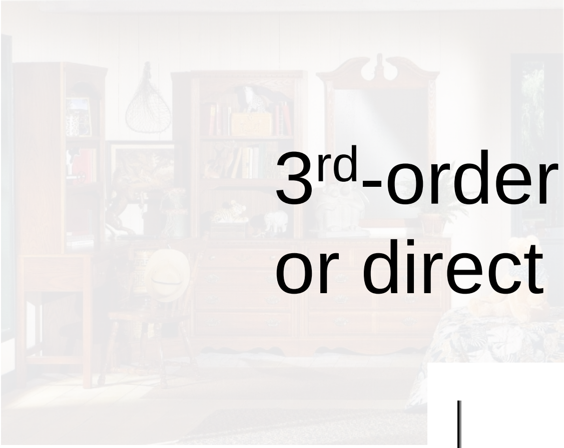
Carpet | channel 260

Floor | channel 459

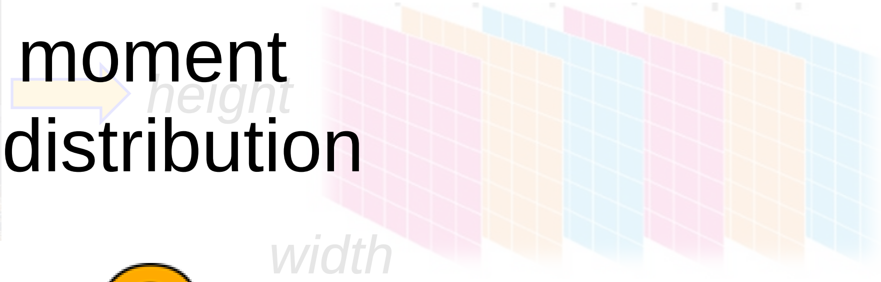
B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. Learning Deep Features for Discriminative Localization. Computer Vision and Pattern Recognition (CVPR), 2016.

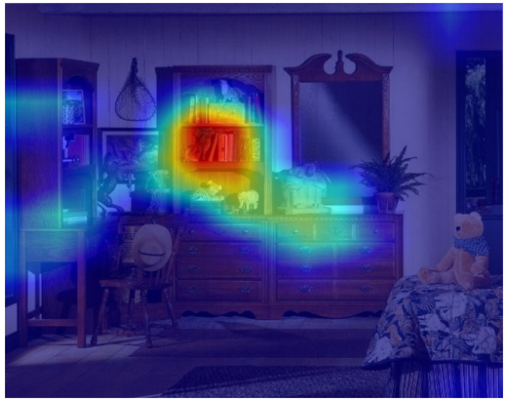
Why Higher-ord

What does each chann



3^rd-order moment
or direct distribution





Bookcase | channel 97

Plant | channel 384

Drawer | channel 360

Decoration | channel 44

Carpet | channel 260

Floor | channel 459

B. Zhou, A. Khosla, … Oliva, and A. Torralba.
Learning Deep Feat… …native Localization. Computer
Vision and Pattern Recognition (CVPR), 2016.

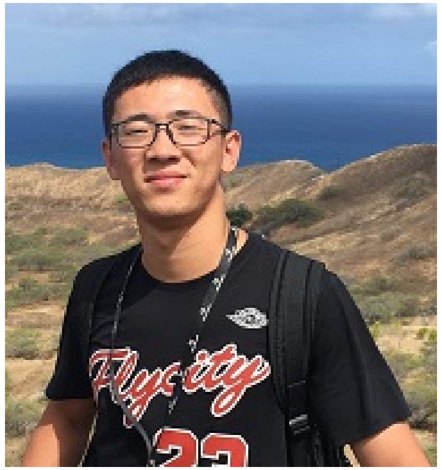
Outline

- What is Higher-order?
- Why We Study High-order
- Overview of Speaker
- Overview of Tutorial

Overview of Speaker

Wangmeng Zuo received the Ph.D. degree in computer application technology from the Harbin Institute of Technology, Harbin, China, in 2007. He is currently a Professor in the School of Computer Science and Technology, Harbin Institute of Technology. His current research interests include image enhancement and restoration, object detection, visual tracking, and image classification. He has published over 70 papers in toptier academic journals and conferences. He has served as a Tutorial Organizer in ECCV 2016, an Associate Editor of the IET Biometrics and Journal of Electronic Imaging, and the Guest Editor of Neurocomputing, Pattern Recognition, IEEE Transactions on Circuits and Systems for Video Technology, and IEEE Transactions on Neural Networks and Learning Systems.

Overview of Speaker



Qilong Wang received the Ph.D. Degree in the School of Information and Communication Engineering, Dalian University of Technology in 2018. He is currently a lecturer in the College of Intelligence and Computing, Tianjin University. His research interests include visual classification and deep probability distribution modeling. He has published several papers in top conferences and referred journals including ICCV, CVPR, ECCV, NIPS, IJCAI, TPAMI, TIP and TCSVT.

Overview of Speaker

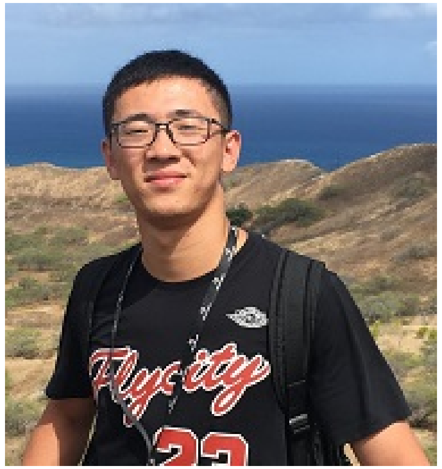
Peihua Li is a professor of Dalian University of Technology. He received Ph.D degree from Harbin Institute of Technology in 2003, and then worked as a postdoctoral fellow at INRIA/IRISA, France. He achieved the honorary nomination of National Excellent Doctoral dissertation in China. He was supported by Program for New Century Excellent Talents in University of Chinese Ministry of Education. His team won 1st place in large-scale iNaturalist Challenge spanning 8000 species at FGVC5 CVPR2018, 2nd place in Alibaba Large-scale Image Search Challenge 2015. His research topics include deep learning and computer vision, focusing on image/video recognition, object detection and semantic segmentation. He has published papers in top journals such as IEEE TPAMI/TIP/TCSVT and top conferences including ICCV/CVPR/ECCV/NIPS. As a principal investigator, he receives funds from National Natural Sceince Foundation of China (NSFC), Chinese Ministry of Education and Huawei Technologies Co., Ltd.

http://peihuali.org/

Outline

- What is Higher-order?
- Why We Study High-order
- Overview of Speaker
- Overview of Tutorial

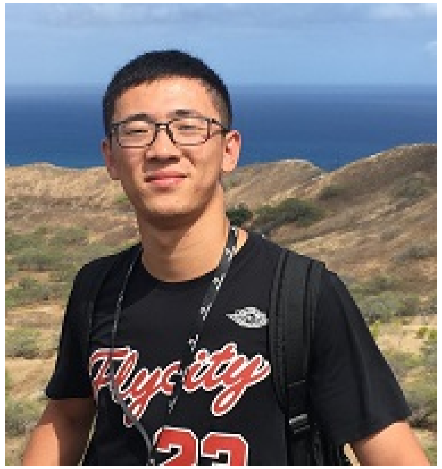
Overview of Tutorial—Part 1



- ➢**Higher-order:** **Yesterday Once More**
- ➢**Higher-order:** **First Dating with CNN**

Overview of Tutorial—Part 2

End-to-End 2nd-order CNN

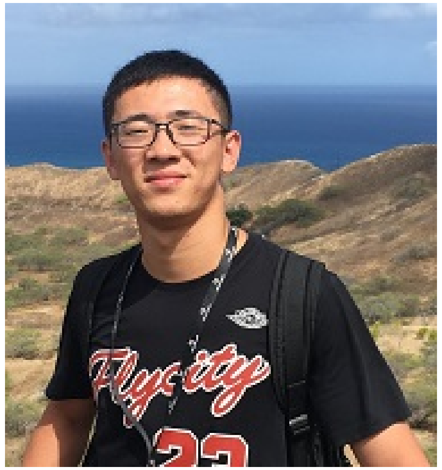


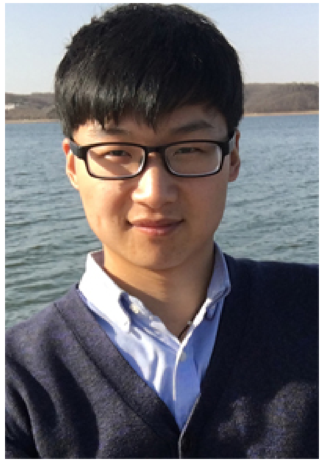
Overview of Tutorial—Part 3

Approximate Higher-order in CNN

Overview of Tutorial—Part 4

Challenge Achievements and Code





Thank you!