

Computer Vision - 2025

Lecture #02. Convolutional Neural Networks

Lectures by Alexei Kornaev ^{1,2,3}

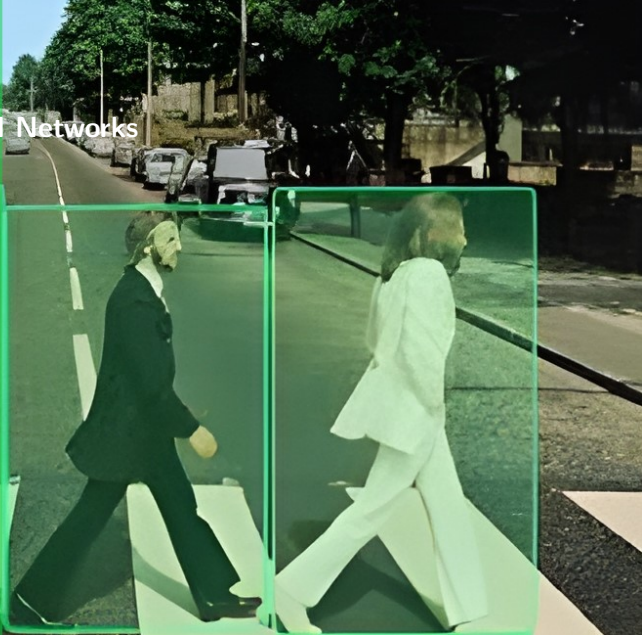
Practical sessions by Kirill Yakovlev ²

¹AI Institute, Innopolis University (IU), Innopolis

²Robotics & CV Master's Program, IU, Innopolis

³RC for AI, National RC for Oncology, Moscow

January 26, 2025



Agenda

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

- ① Outcomes
- ② Key Ideas in CNNs
- ③ CNNs: the Current State and Prospects
- ④ Conclusion

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

Section 1. Outcomes

Outcomes

CV-2025

A.Kor-naev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

This week's lecture and practice on CNNs is designed to provide a deep understanding of the theoretical concepts and practical skills necessary for working with CNN-based architectures. By the end of this lecture, students will be able to:

- 1 Understand core concepts: convolutions, filters, feature maps, padding, pooling, and key CNN architectures (e.g., ResNet, DenseNet, ConvNeXt).
- 2 Apply transfer learning and fine-tuning to adapt pre-trained models for new tasks using tools like PyTorch.
- 3 Implement CNNs on datasets like MNIST and CIFAR-10, bridging the gap from educational tasks to real-world applications.

Key Takeaway: Transfer learning accelerates solving practical problems by leveraging pre-trained models.

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

Section 2. Key Ideas in CNNs

Physiology of cats

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

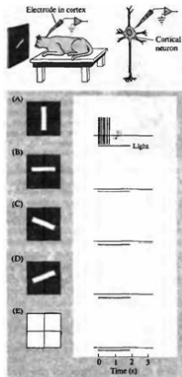


Figure: Responses of the cat's visual cortex cell.

Implementation in math.

The horizontal derivative kernel approximates $\frac{\partial I}{\partial x}$ using **finite differences**. For images, this translates to computing intensity changes along the x-axis. The Sobel kernel for horizontal derivative is:

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

This kernel computes:

$$\frac{\partial I}{\partial x} \approx (I(x+1, y) - I(x-1, y))$$

and incorporates smoothing to reduce noise. It detects horizontal edges by highlighting intensity changes from left to right.

Hands-on coding: **Kernels** [CV-2025]

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

#	Filter	Description
1	Median Filtering	Replaces each pixel with the median value of its neighborhood.
2	Gaussian Blur	A Gaussian kernel. Reduces noise but blurs edges.
3	Bilateral Filtering	Smooths the image while preserving edges by considering both spatial and intensity differences.
4	Non-Local Means Denoising	Removes noise by averaging similar patches across the image. Computationally intensive.
5	Histogram Equalization	Redistributes pixel intensities to improve contrast. Simple and effective for enhancing details.
6	Edge Detection (Canny)	Detects edges using gradient-based methods and hysteresis thresholding. Preserves strong edges.
7	Thresholding	Converts a grayscale image to binary based on a threshold value. Simple pixel-wise operation.
8	Morphological Operations (Closing)	Closes small defects in binary images using a structuring element.
9	Inpainting	Fills missing or corrupted regions in the image using surrounding information. Solves partial differential equations.

Table: Filters.

Padding, Pooling, and Striding

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

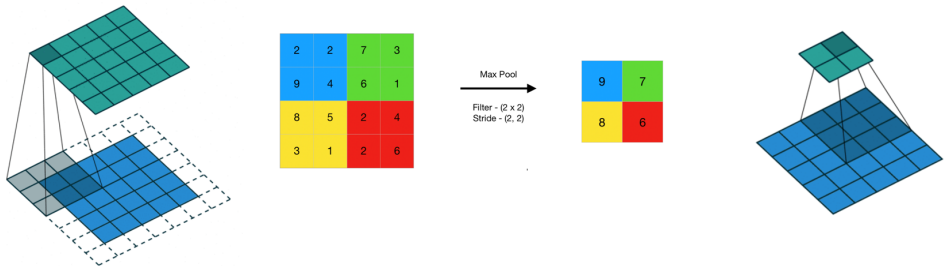


Figure: Padding (left), Pooling (middle), and Striding (right). CNNs by Neurohive.

Padding, Pooling, and Striding

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

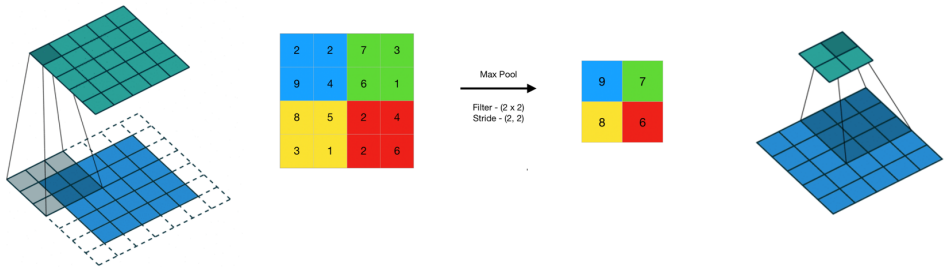


Figure: Padding (left), Pooling (middle), and Striding (right). CNNs by Neurohive.

2D Convolution Applied to an Image

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

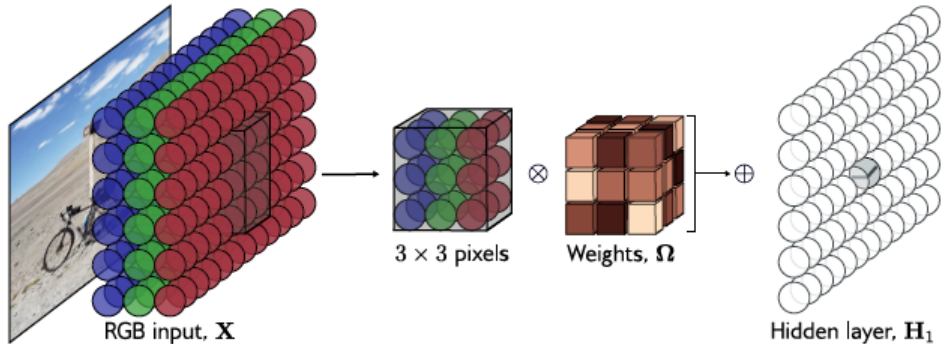


Figure: The image is treated as a 2D input with three channels corresponding to the red, green, and blue components. With a 3×3 kernel, each pre-activation in the first hidden layer is computed by pointwise multiplying the $3 \times 3 \times 3$ kernel weights with the $3 \times 3 \times 3$ RGB image patch centered at the same position, summing, and adding the bias. To calculate all the pre-activations in the hidden layer, we “slide” the kernel over the image in both horizontal and vertical directions. The output is a 2D layer of hidden units. To create multiple output channels, we would repeat this process with multiple kernels, resulting in a 3D tensor of hidden units at hidden layer \mathbf{H}_1 [Prince, 2023]

Hands-on the book by Howard and Gugger [2020]

CV-2025

A.Kornaev,
K.Yakovlev

CNN utilizes the features of the visual cortex, where simple cells are activated by simple features (such as lines), and complex cells by combinations of activations of simple cells. The CNN is associated with the mathematical operation of convolution for reducing matrix sizes.

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

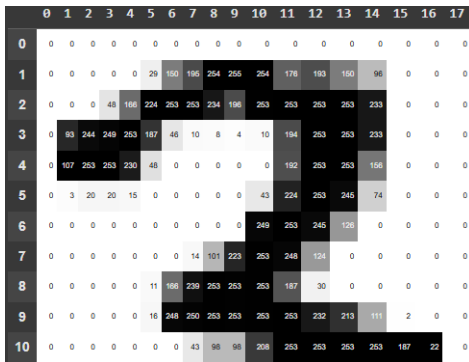


Figure: Check the code: 13 convolutions.ipynb, and 14 resnet.ipynb

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

Section 3. CNNs: the Current State and Prospects

Paper Reading

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

The following review papers on CNNs [Kiranyaz et al., 2021; Li et al., 2021] are the most cited in 2025:

- ① 1D Convolutional Neural Networks and Applications - A Survey
- ② A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

Section 4. Conclusion

Conclusion

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in
CNNs

CNNs: the
Current State
and Prospects

Conclusion

Summary:

- Images in computer vision are numerical representations of visual data, stored as pixel grids.
- Different types of images (grayscale, color, binary, multispectral) serve various purposes in CV applications.
- Key attributes like resolution, bit depth, and storage formats play critical roles in image processing.

Key Takeaway:

- Understanding image structures and properties is fundamental to developing effective CV models and applications.

Bibliography

CV-2025

A.Kornaev,
K.Yakovlev

Outcomes

Key Ideas in CNNs

CNNs: the Current State and Prospects

Conclusion

CV-2025. The course repository. <https://github.com/Mechanics-Mechatronics-and-Robotics/CV-2025>, 2025. Accessed: January 26, 2025.

J. Howard and S. Gugger. **Deep Learning for Coders with Fastai and Pytorch: AI Applications Without a PhD**. O'Reilly Media, Incorporated, 2020. ISBN 9781492045526. URL <https://books.google.no/books?id=xd6LxgEACAAJ>.

Serkan Kiranyaz, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, and Daniel J Inman. 1d convolutional neural networks and applications: A survey. **Mechanical systems and signal processing**, 151:107398, 2021.

Zewen Li, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. A survey of convolutional neural networks: analysis, applications, and prospects. **IEEE transactions on neural networks and learning systems**, 33(12):6999–7019, 2021.

Simon J.D. Prince. **Understanding Deep Learning**. The MIT Press, 2023. URL <http://udlbook.com>.