

# Stanford Univ. CS231n 스터디 2팀

2025.04.29

발표자 : 양재영


## 스터디원 소개 및 만남 인증

방세현(20230674)

방세현(20230674)

최재민

최재민



양재영

이가연

이가연

1 건

1 스터디룸02

2025-03-20 17:00 ~ 19:00

유형 : 스터디룸 (서울)

위치 : 서울학술정보원 학술정보원

동반자수 : 3

1 스터디룸03

2025-03-27 17:00 ~ 19:00

유형 : 스터디룸 (서울)

위치 : 서울학술정보원 학술정보원

동반자수 : 3

1 스터디룸02

2025-04-03 17:00 ~ 18:00

유형 : 스터디룸 (서울)

위치 : 서울학술정보원 학술정보원

동반자수 : 3

스터디원 1 : 방세현

스터디원 2 : 이가연

스터디원 3 : 양재영



스터디원 4 : 최재민


# 목차

1. 스터디 주제
2. 스터디 진행 방식
3. 스터디 진행 현황
4. 스터디 내용 공유

# 스터디 주제

[CS231n Home](#) [Schedule](#) [Assignments](#) [Project](#) [Office Hours](#) [Lecture Videos](#) [Ed](#) [Useful Notes](#)

 CS231n: Deep Learning for Computer Vision   
Stanford - Spring 2025



\*This network is running live in your browser

### Course Description

Computer Vision has become ubiquitous in our society, with applications in search, image understanding, apps, mapping, medicine, drones, and self-driving cars. Core to many of these applications are visual recognition tasks such as image classification, localization and detection. Recent developments in neural

# 스터디 진행 방식

- 대면 스터디
- 매주 유튜브에 업로드 된 2017년 강의 시청 후 발제
- 2025년 버전에 추가된 내용 또한 추가적으로 공부 후 공유

Date	Lecture
3/13	OT
3/20	Introduction(Lecture 1), Deep learning Basics(Lecture 2, 3)
3/27	Deep learning Basics(Lecture 4), Perceiving and Understanding the Visual World(Lecture 5, 6)
4/3	Perceiving and Understanding the Visual World(Lecture 7, 8, 9)
5/1	Perceiving and Understanding the Visual World(Lecture 10, 11), Generative and Interactive Visual Intelligence(Lecture 12)
5/8	Generative and Interactive Visual Intelligence(13, 14)
5/15	Generative and Interactive Visual Intelligence(15, 16)
5/22	Generative and Interactive Visual Intelligence(17, 18)

# 스터디 진행 현황

CS231n Study

4월 3일 편집 공유

## CS231n Study

표

### Lecture 정리

Aa 이름	사람	
Lecture3. Loss Function and Optimization	J_young	
Lecture4. Introduction to Neural Networks	Choi Jaemin	
Lecture 5	Igy7721@cau.ac.kr	
Lecture6. Training Neural Networks 1. (Activation function)	SeHyun Bang	
Lecture7. Training Neural Networks 2	J_young	
Lecture8. Deep Learning Software	Choi Jaemin	

+ 새 페이지

### 캘린더 보기

2025년 4월

일	월	화	수	목	금	토
30	31	4월 1일	2	3	4	5
				Meetings		

?

# 스터디 진행 현황

## CS231n Study

표

### Lecture 정리

Aa 이름

사람

Lecture3. Loss Function and Optimization

J\_young

Lecture4. Introduction to Neural Networks

Choi Jaemin

Lecture 5

lgy7721@cau.ac.kr

Lecture6. Training Neural Networks 1. (Activation function)

SeHyun Bang

Lecture7. Training Neural Networks 2

J\_young

Lecture8. Deep Learning Software

Choi Jaemin

새 페이지

캘린더 보기

2025년 4월

< 오늘 >

일

월

화

수

목

금

토

30

31

4월 1일

2

3

4

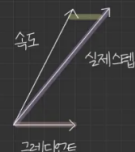
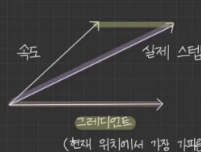
5

## Nesterov Momentum

$$V_{t+1} = \rho V_t + \nabla f(x_t)$$

$$x_{t+1} = x_t - \alpha V_{t+1}$$

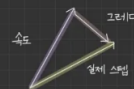
오버슈팅 문제



Momentum의 단점 중 하나는 Overshooting임.

이를 해결하기 위해 등장한 것이 Nesterov Momentum!

## Nesterov Momentum



· 속도 벡터 : 현재 이동하는 속도

· 그래디언트 벡터

: 현재 속도로 한 걸음 미리 가본 위치에서 내려막을 방향을 의미

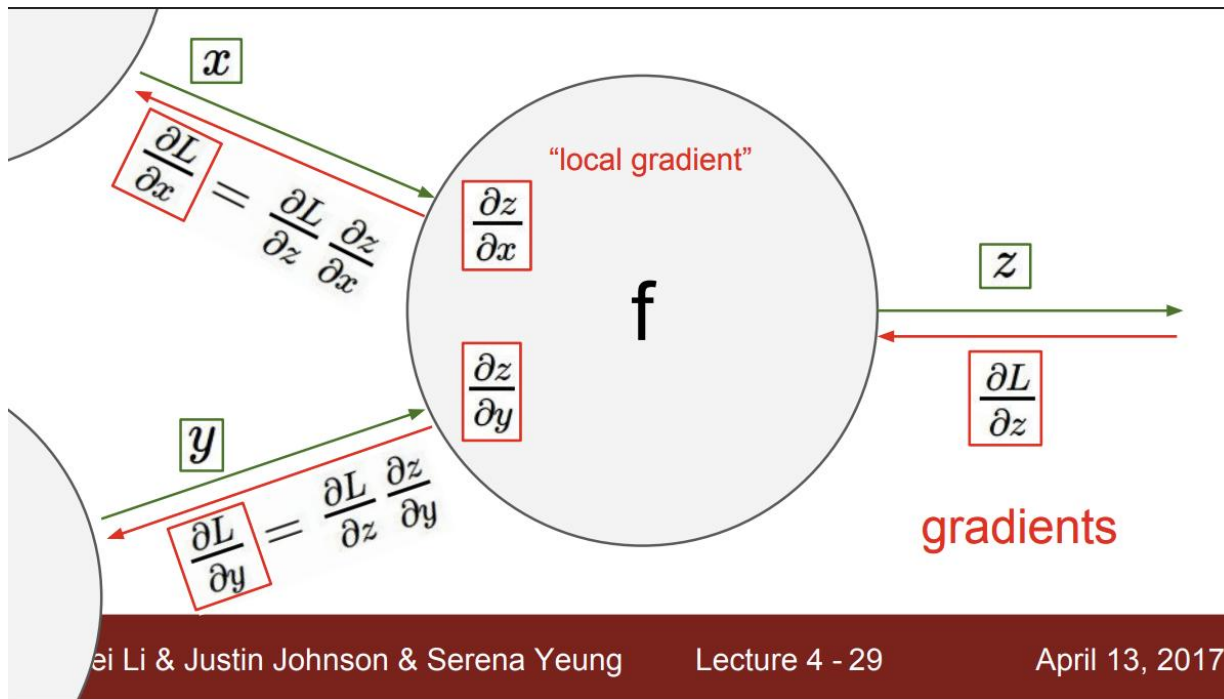
→ 한 걸음 미리 가기

$$V_{t+1} = \rho V_t - \alpha \nabla f(x_t + \rho V_t)$$

$\rho$  : 미량계수

# 스터디 내용 공유

## Lecture\_4. Introduction to Neural Networks





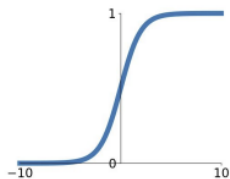
# 스터디 내용 공유

## Lecture\_4. Introduction to Neural Networks

### Activation functions

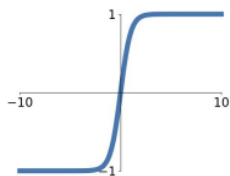
#### Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



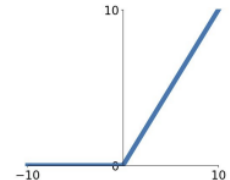
#### tanh

$$\tanh(x)$$



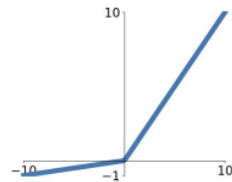
#### ReLU

$$\max(0, x)$$



#### Leaky ReLU

$$\max(0.1x, x)$$

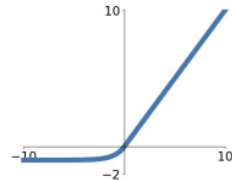


#### Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

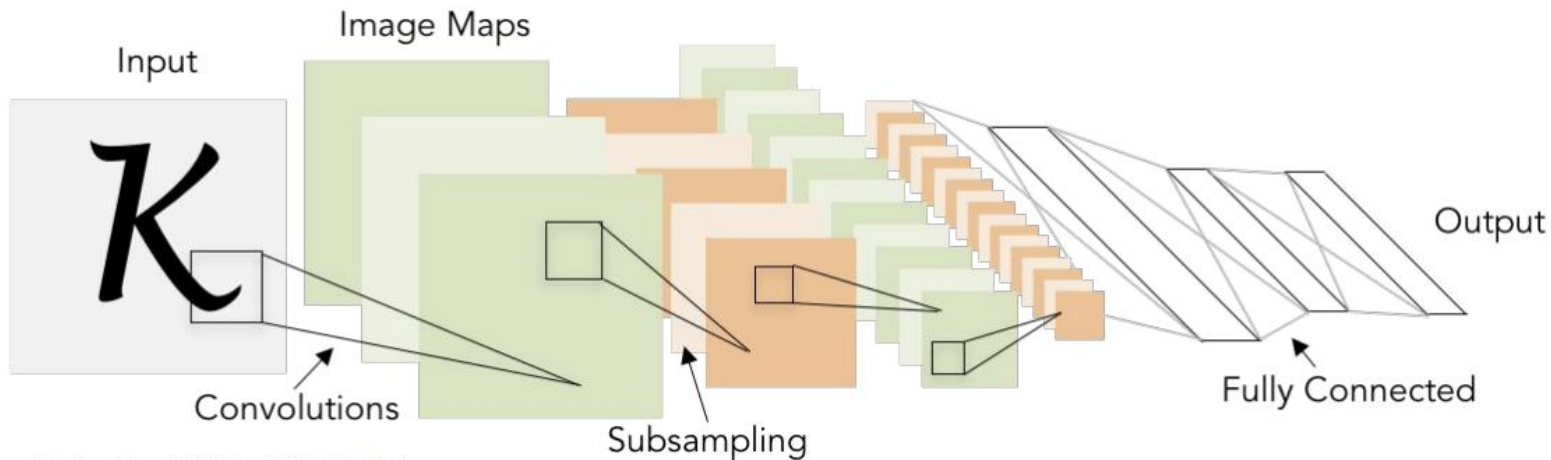
#### ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



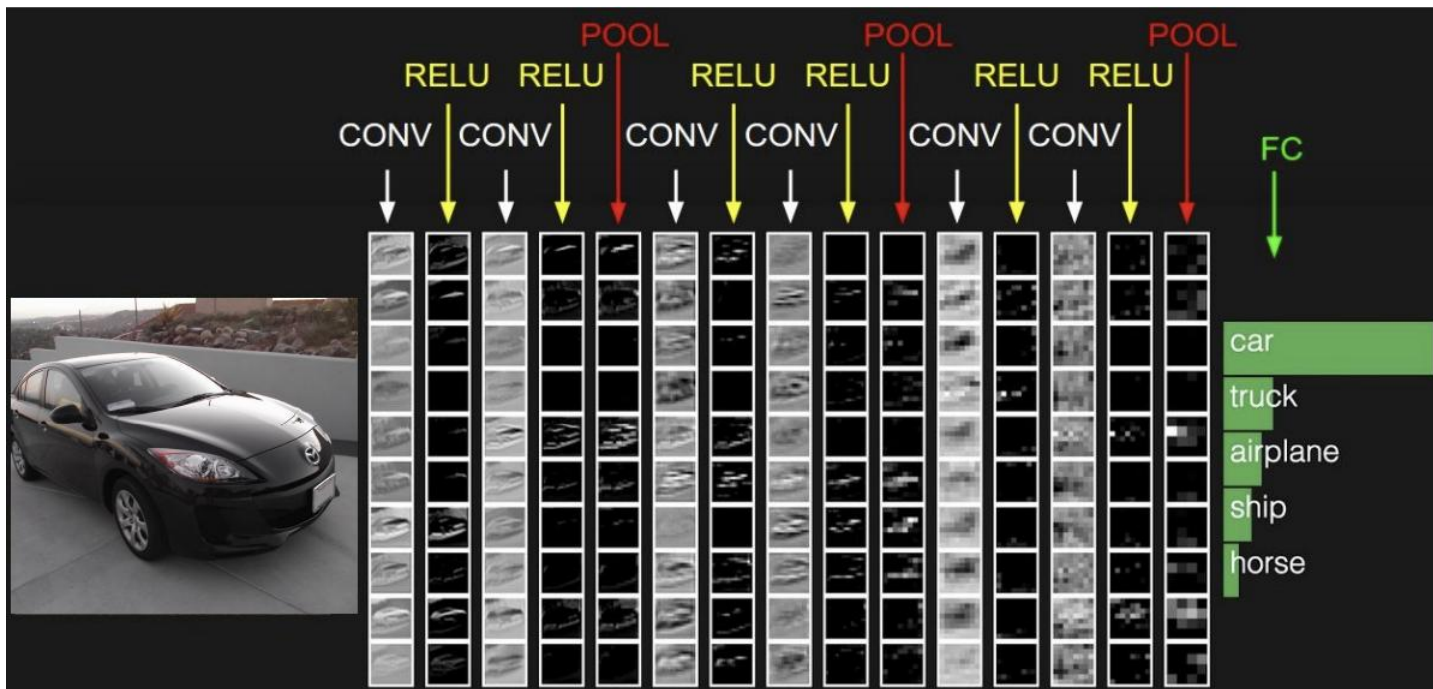
# 스터디 내용 공유

## Lecture\_5. Convolutional Neural Networks



# 스터디 내용 공유

## Lecture\_5. Convolutional Neural Networks

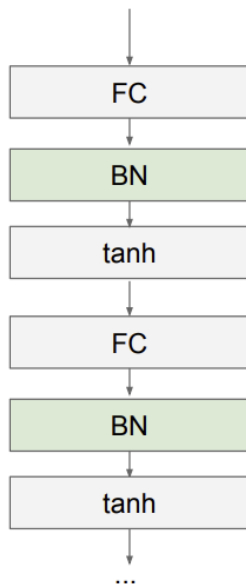


# 스터디 내용 공유

## Lecture\_6-7. Training Neural Networks

### Batch Normalization

[Ioffe and Szegedy, 2015]

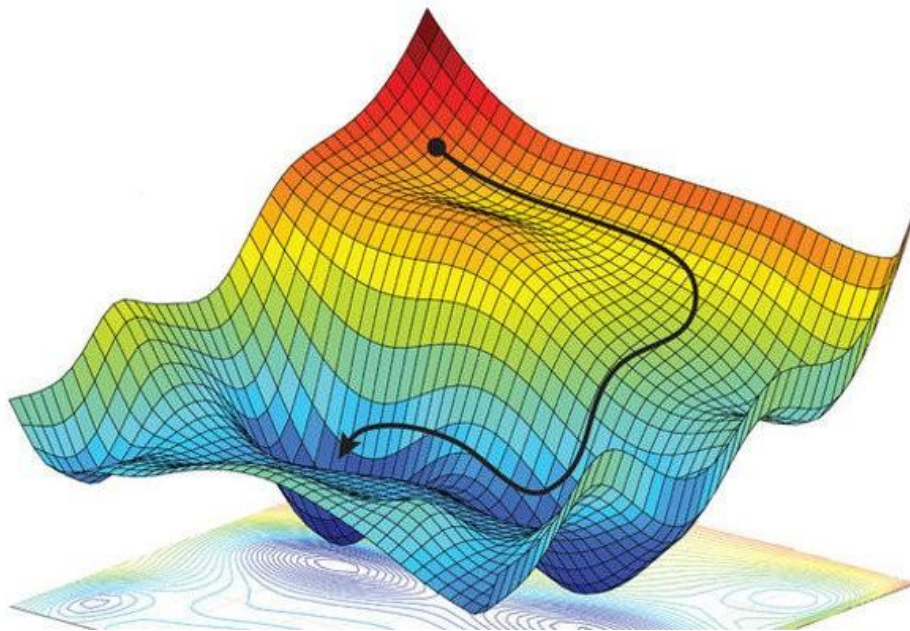


Usually inserted after Fully Connected or Convolutional layers, and before nonlinearity.

$$\hat{x}^{(k)} = \frac{x^{(k)} - E[x^{(k)}]}{\sqrt{\text{Var}[x^{(k)}]}}$$

# 스터디 내용 공유

## Lecture\_6-7. Training Neural Networks



Gradient Descent  
SGD  
Nesterov Momentum  
RMS Prop  
ADAM

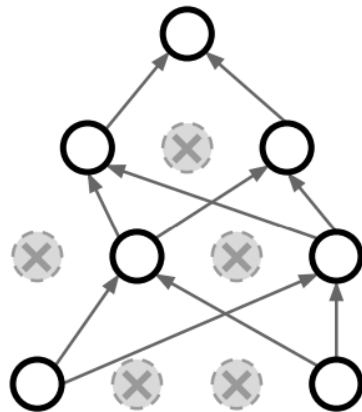
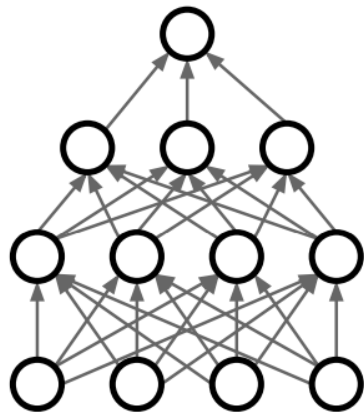
.  
. .  
. .

# 스터디 내용 공유

## Lecture\_6-7. Training Neural Networks

### Regularization: Dropout

In each forward pass, randomly set some neurons to zero  
Probability of dropping is a hyperparameter; 0.5 is common



감사합니다