Үйірткілі нейрондық желілер (CNN) және компьютерлік көру

Кескіннің кеңістіктік құрылымын нейрондық желі қалай пайдаланады?

Дәрістің негізгі міндеттері

- Үйірткі, ядро және белгілер картасының мағынасын түсіндіру.
- Stride, padding, pooling және Flatten амалдарын ажырату.
- CNN архитектурасының өлшемдері мен параметрлерін есептеу.

Кескінді ұсыну және MLP шектеулері

Кескін — сандық массив

Сұр реңкті: H × W
RGB: H × W × 3

224 × 224 × 3 = 150 528 мән.

Көршілес пиксельдердің орналасуы маңызды.

Dense қабатындағы шығын

28 × 28 → 1000 нейрон:
784 000 салмақ.

224 × 224 × 3 → 1000 нейрон:
150 528 000 салмақ.

Мұнда bias есептелмеген.

CNN жергілікті байланыстарды және салмақтарды ортақ пайдалануды енгізеді.

Үйірткі операциясы және белгілер картасы

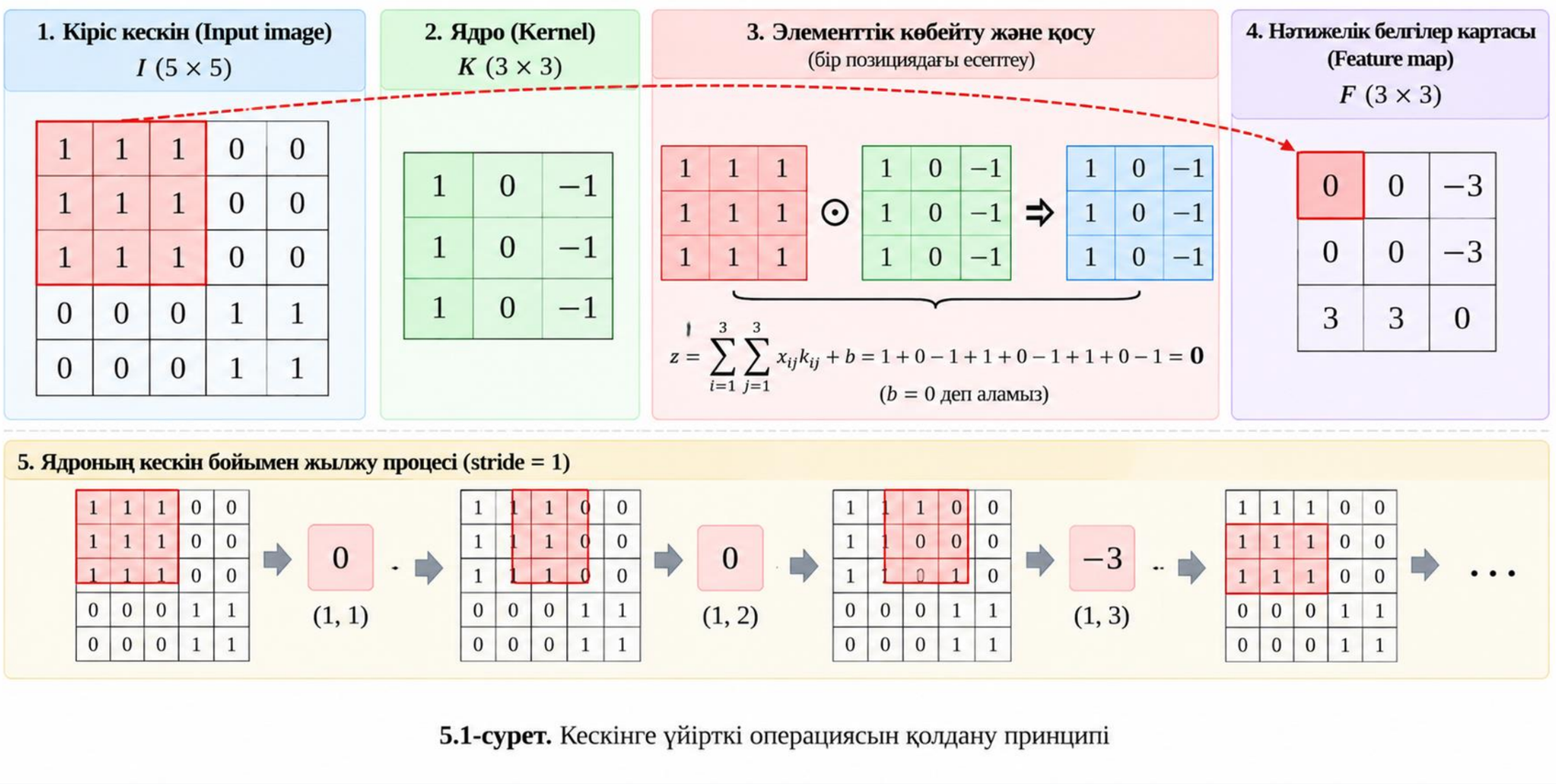
$$z = \Sigma_i \Sigma_j x_{ij} k_{ij} + b$$

Kernel — кескіннің шағын аймағын өңдейтін салмақтар жиыны.
Бір позицияда сәйкес элементтер көбейтіліп, қосылады.
Ядро жылжығанда нәтижелер белгілер картасын құрайды.

Келесі қадам: активацияны қолдану, мысалы A = ReLU(Z).

DL-дегі «convolution» әдетте ядро аударылмайтын кросс-корреляция түрінде есептеледі.

Үйірткі принципі: дәрістегі схема



5.1-сурет. Кескінге үйірткі операциясын қолдану принципі

Ескерту: суреттегі нәтижелік матрицада есептік сәйкессіздік бар.
Берілген I және K үшін тексерілген нәтиже келесі слайдта көрсетілген.

Дереккөз: 5-дәріс конспектісі, 2-бет, 5.1-сурет. Түпнұсқа сақталған.

5.1-суреттегі мысалдың тексерілген нәтижесі

Суреттегі ядро K

1	0	-1
1	0	-1
1	0	-1

Дұрыс белгілер картасы Z

0	3	3
0	1	1
0	-1	-1

Мысалы, жоғарғы жолдың 2-элементі:
$(1 \times 1 + 1 \times 0 + 0 \times (-1)) \times 3 = 3$.

Шарттар: stride = 1, padding = 0, bias = 0; активация әлі қолданылған жоқ.

Ортақ параметрлер және бірнеше сүзгі

Parameter sharing

Бір ядроның коэффициенттері барлық кеңістіктік позицияда бірдей.

Бір арналы 3 × 3 сүзгі:
9 салмақ + 1 bias = 10 параметр.

Бірнеше арна және сүзгі

RGB кірісінде бір сүзгінің өлшемі:
3 × 3 × 3.

32 сүзгі → 32 белгілер картасы.
Әр сүзгі барлық кіріс арнасын өңдейді.

Ядро коэффициенттері оқыту барысында анықталады; олар міндетті түрде қолмен берілмейді.

Stride және padding: шығыс өлшемі

$$H_{шығ} = \lfloor (H + 2P - K) / S \rfloor + 1$$

S — ядроның қозғалу қадамы. P — әр жақтағы толықтыру.
Ені W үшін де осы ереже қолданылады.

Кіріс H	K	S	P	Шығыс H
5	3	1	0	3
5	3	1	1	5

Қолданылу шарты: dilation = 1, симметриялы padding.

ReLU және Max Pooling

ReLU

$A = \max(0, Z)$

Теріс мәндер нөлге ауысады.

Бейсызық тәуелділіктерді модельдеуге мүмкіндік береді.

Оқытылатын параметрі жоқ.

Max Pooling

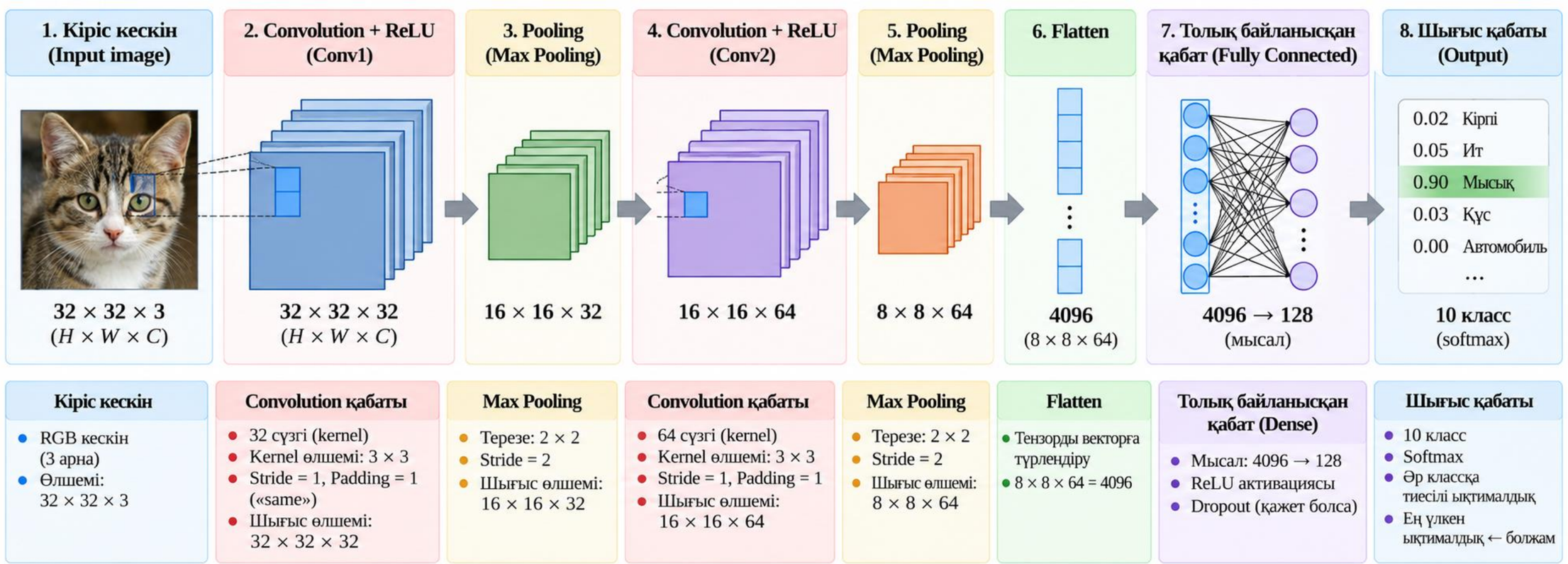
$\max([[1, 5], [2, 3]]) = 5$

2 × 2 терезе, stride = 2:
28 × 28 → 14 × 14.

Кеңістіктік өлшем азаяды.
Оқытылатын параметрі жоқ.

Pooling тек өлшемді азайтпайды: жергілікті мәндердің бір бөлігі де жоғалады.

CNN архитектурасының жалпы құрылымы



5.2-сурет. Үйірткілі нейрондық желінің (CNN) жалпы архитектурасы

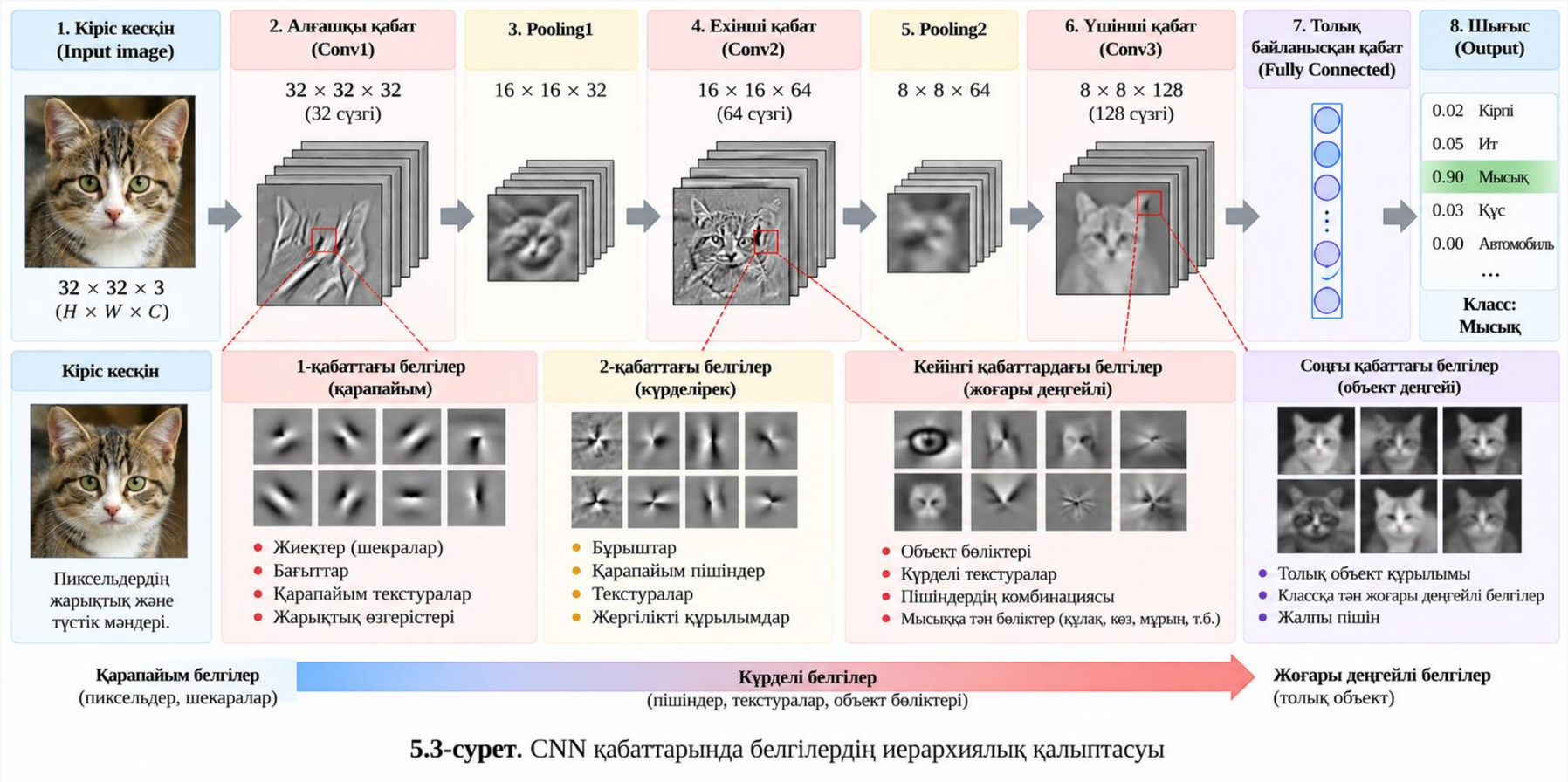
Дереккөз: 5-дәріс конспектісі, 5-бет, 5.2-сурет. Түпнұсқа.

Архитектурадағы өлшемдер және Flatten

Кезең	Шығыс өлшемі	Негізгі өзгеріс
Кіріс	32 × 32 × 3	RGB кескін
Conv + ReLU	32 × 32 × 32	32 белгілер картасы
Pooling	16 × 16 × 32	Кеңістік азаяды
Conv + ReLU	16 × 16 × 64	64 белгілер картасы
Pooling	8 × 8 × 64	Кеңістік азаяды
Flatten	4096	Пішін ғана өзгереді

Flatten: 8 × 8 × 64 = 4096; оқытылатын параметрлер жоқ.

Белгілердің иерархиялық қалыптасуы



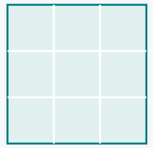
5.3-сурет. CNN қабаттарында белгілердің иерархиялық қалыптасуы

Схема — түсіндірмелік үлгі; әр нейронға нақты объект бөлігін телу міндетті емес.

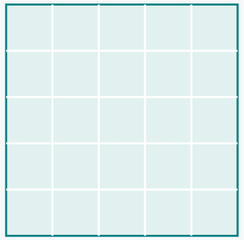
Дереккөз: 5-дәріс конспектісі, 6-бет, 5.3-сурет. Түпнұсқа.

Қабылдау өрісі — Receptive Field

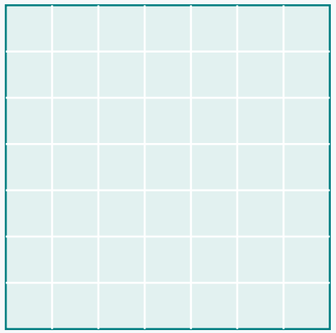
Қабылдау өрісі — кірістің берілген нейрон мәніне әсер ететін аймағы.



1-қабат: 3 × 3



2-қабат: 5 × 5



3-қабат: 7 × 7

Мысал: 3 × 3 ядролар, stride = 1, pooling жоқ; теориялық өріс қабат сайын өседі.

Convolution қабатындағы параметрлер саны

$$N = K_h \times K_w \times C_{in} \times C_{out} + C_{out}$$

Ядро: 3 × 3; кіріс: 3 арна; шығыс: 32 сүзгі.

$$N = 3 \times 3 \times 3 \times 32 + 32 = 896$$

Параметр саны H пен W-ге тәуелсіз; есептеу көлемі мен жад қажеттілігі оларға тәуелді.

CNN қалай оқытылады және класс қалай таңдалады?

Кескін → Forward → Loss → Backpropagation → Жаңарту

$$K \leftarrow K - \eta\, \partial J/\partial K$$

Үйірткі ядролары, Dense салмақтары және ығысулар жаңартылады.

$$\hat{y} = (\hat{y}_1, \ldots, \hat{y}_{10}),\ \hat{c} = \arg\max_i \hat{y}_i$$

10 кластың бірін таңдау: ең үлкен шығыс бағасына сәйкес класс.

CNN мүмкіндіктері, шектеулері және келесі есеп

Артықшылықтары

Жергілікті кеңістіктік байланыстар.
Параметрлерді ортақ пайдалану.
Белгілерді автоматты үйрену.
Иерархиялық көріністер қалыптастыру.

Шектеулері

Деректер мен есептеу ресурстарына сұраныс.
Белгілерді түсіндірудің қиындығы.

Кәдімгі жіктеу шығысы объектінің орнын көрсетпейді.

Classification: «Не бар?» Detection / YOLO: «Не бар және қай жерде?»

Бақылау сұрақтары

01 MLP-мен салыстырғанда CNN кескін құрылымын қалай пайдаланады?

02 Kernel, feature map және parameter sharing қалай байланысады?

03 Stride пен padding шығыс өлшеміне қалай әсер етеді?

04 ReLU, pooling және Flatten қызметтері қалай ажыратылады?

05 Қабылдау өрісі мен параметрлер саны қалай анықталады?

06 Кескінді жіктеу мен нысандарды анықтаудың айырмасы қандай?

Дереккөз және сандық нақтылау

Негізгі дереккөз

«5-дәріс. Үйірткілі нейрондық желілер (CNN) және компьютерлік көру». Дәріс конспектісі, 9 бет.

Түпнұсқа иллюстрациялар

5.1-сурет: үйірткі операциясы, 2-бет → 4-слайд.
5.2-сурет: CNN архитектурасы, 5-бет → 9-слайд.
5.3-сурет: белгілер иерархиясы, 6-бет → 11-слайд.

5.1-суреттегі нәтижелік матрица қайта есептелді: дұрыс мәндер 5-слайдта берілген. Қабылдау өрісінің 3 → 5 → 7 мысалы және кросс-корреляция туралы түсіндірме — қысқа әдістемелік нақтылаулар.