Нысандарды анықтауға арналған нейрондық желілер. YOLO архитектурасы

Кескінде не бар және ол қай жерде орналасқан?

Дәрістің негізгі міндеттері

- Bounding box, confidence, IoU және NMS ұғымдарын ажырату.
- Backbone, Neck және Detection Head қызметтерін түсіндіру.
- Детекторды оқыту және бағалау қағидаларын талдау.

Жіктеу, локализация және нысандарды анықтау

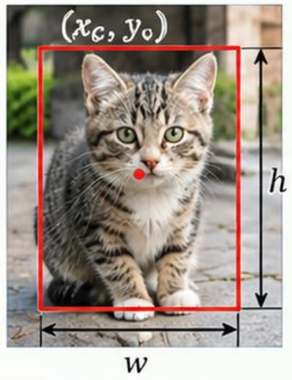
1. Кескінді классификациялау (Image Classification)



Мысық (Cat)

- **Кіріс:** бүкіл кескін
- **Шығыс:** бір класс
- **Мысал:** «Мысық»
- Нысанның орны анықталмайды

2. Нысанды локализациялау (Object Localization)



Мысық (Cat)

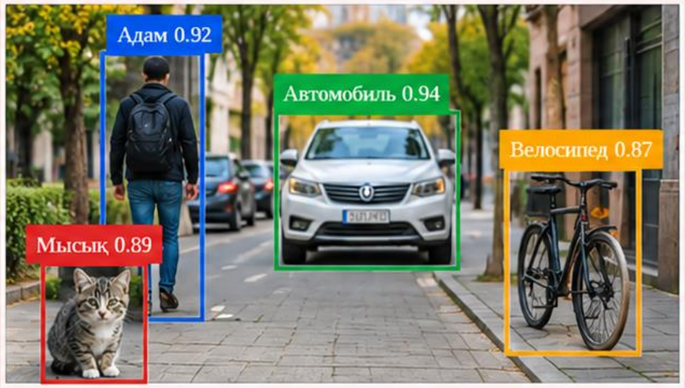
- **Кіріс:** бүкіл кескін
- **Шығыс:** класс + bounding box
- Bounding box параметрлері:

$$B = (x_c, y_c, w, h)$$

x_c, y_c – орталық координаталары
w – ені
h – биіктігі

- **Мысал:** «Мысық» + bounding box

3. Нысандарды анықтау (Object Detection)



- **Кіріс:** бүкіл кескін
- **Шығыс:** бірнеше нысан, әрқайсысы үшін:
  - класс (мысалы, адам, автомобиль, велесипед, мысық)
  - bounding box (x_c, y_c, w, h)
  - сенімділік мәні (confidence score)
- Нысанның класы және орны бір уақытта анықталады

6.1-сурет. Image Classification, Localization және Object Detection есептерінің айырмашылығы

Дереккөз: 6-дәріс конспектісі, 2-бет, 6.1-сурет.

Bounding box: орын және өлшем

$$B = (x_c, y_c, w, h)$$

Орталық координаталары

x_c = 0,5: кескін енінің 50%-ы.
y_c = 0,4: кескін биіктігінің 40%-ы.

Шекаралық қорап өлшемі

w = 0,3: кескін енінің 30%-ы.
h = 0,6: кескін биіктігінің 60%-ы.

Нормаланған координаталарда x пен w — енге, y пен h — биіктікке қатысты.

Детектор шығысы және confidence score

Әр болжам: bounding box + класс бағалары + сенімділік ұпайы.

Confidence threshold

Мысал: шек = 0,5.

Score = 0,95 → сақталады.
Score = 0,18 → алынып тасталады.

Маңызды айырмашылық

Confidence — модельдің болжам ұпайы.

IoU — екі қораптың геометриялық сәйкесу өлшемі.
Бұл екеуі бір көрсеткіш емес.

Score құрамының нақты анықтамасы YOLO нұсқасы мен іске асыруына тәуелді.

Intersection over Union — IoU

$$\text{IoU} = \text{Ақиылыс} / \text{Абірігу}$$

$$\text{Абірігу} = A_1 + A_2 - \text{Ақиылыс}$$

IoU = 0: қиылысу жоқ. IoU = 1: қораптар толық сәйкес.
Мысал: Ақиылыс = 60, Абірігу = 100 → IoU = 0,6.

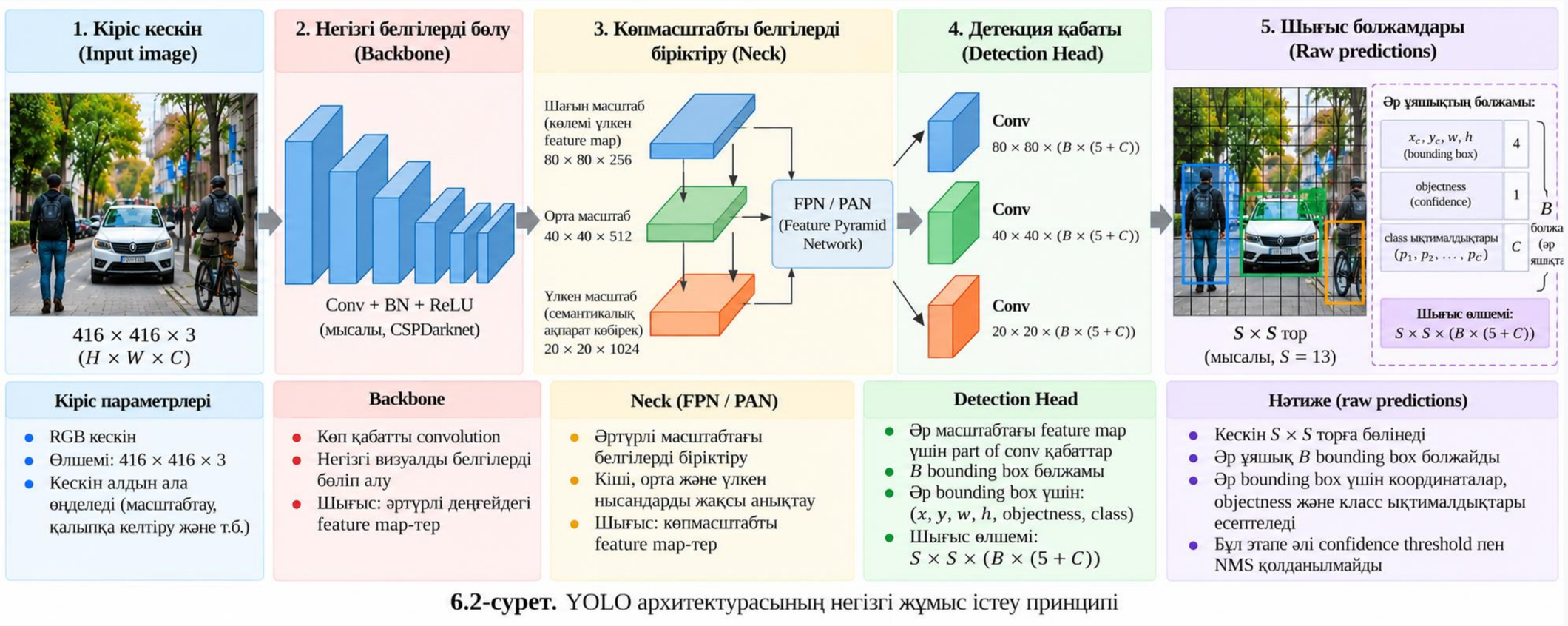
Бағалауда: болжам ↔ нақты қорап. NMS кезінде: болжам ↔ болжам.

Non-Maximum Suppression — NMS

1. Реттеу	Болжамдарды confidence score бойынша кему ретімен қарастыру.
2. Таңдау	Ең жоғары ұпайлы қорапты сақтау.
3. Салыстыру	Сол кластың қалған қораптарымен IoU есептеу.
4. Қайталау	IoU шегінен асатындарды алып тастап, қалғандармен жалғастыру.

Стандартты NMS қораптарды сүзеді; сақталған болжамның score мәнін өзгертпейді.

YOLO архитектурасының негізгі бөліктері



6.2-сурет. YOLO архитектурасының негізгі жұмыс істеу принципі

Схема — тұжырымдамалық үлгі; көрсетілген өлшемдер нақты бір нұсқаның бірыңғай есебі емес.

Дереккөз: 6-дәріс конспектісі, 4-бет, 6.2-сурет.

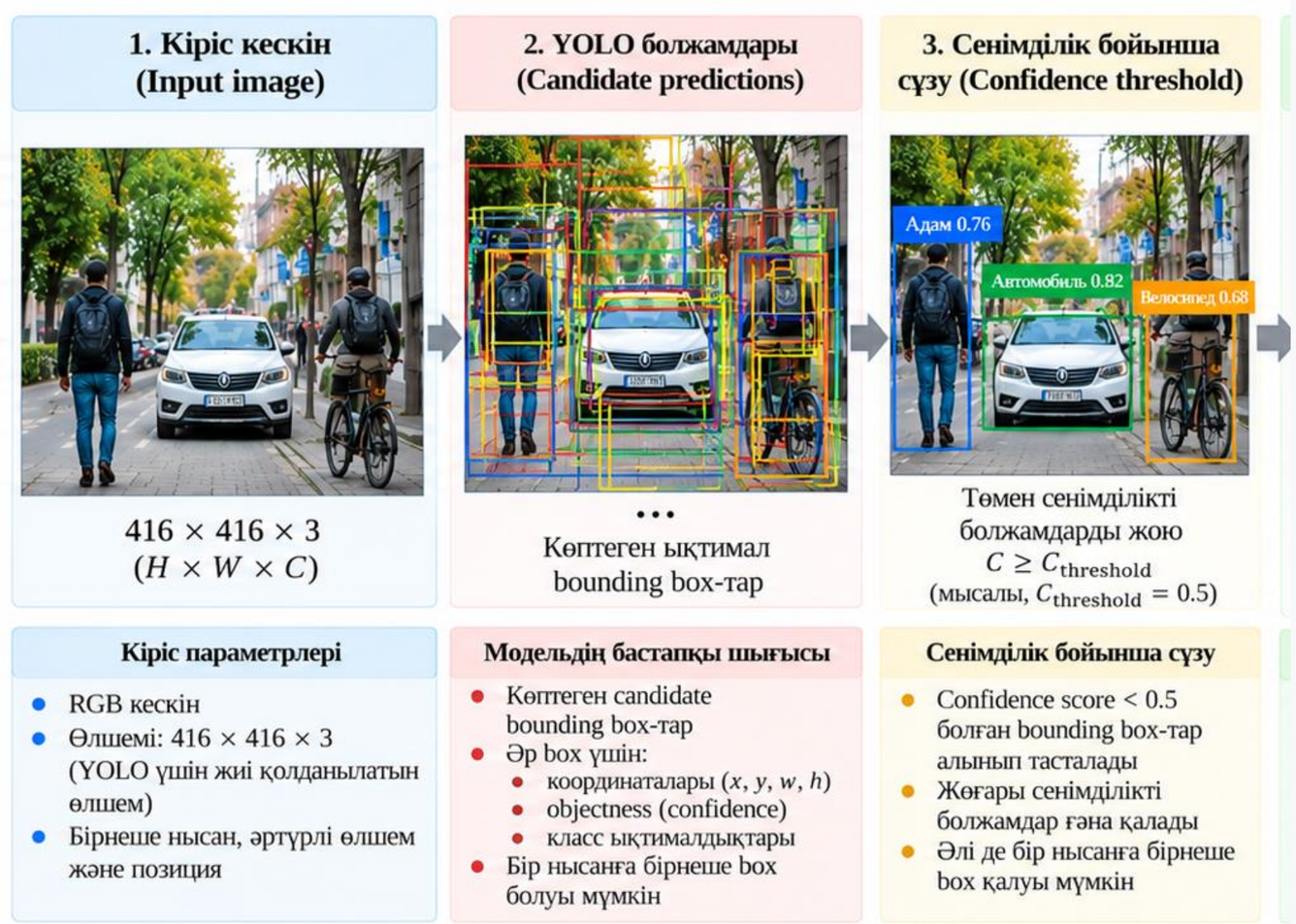
Backbone, Neck, Head және көпмасштабтылық

Бөлік	Қызметі
Backbone	Кескіннен әртүрлі деңгейдегі визуалды белгілерді алу.
Neck	Әртүрлі масштабтағы белгілерді біріктіру.
Detection Head	Қораптар мен класс ұпайларын болжау.

Stride 8 / 16 / 32 мысалында:
640 × 640 → 80 / 40 / 20; 416 × 416 → 52 / 26 / 13.

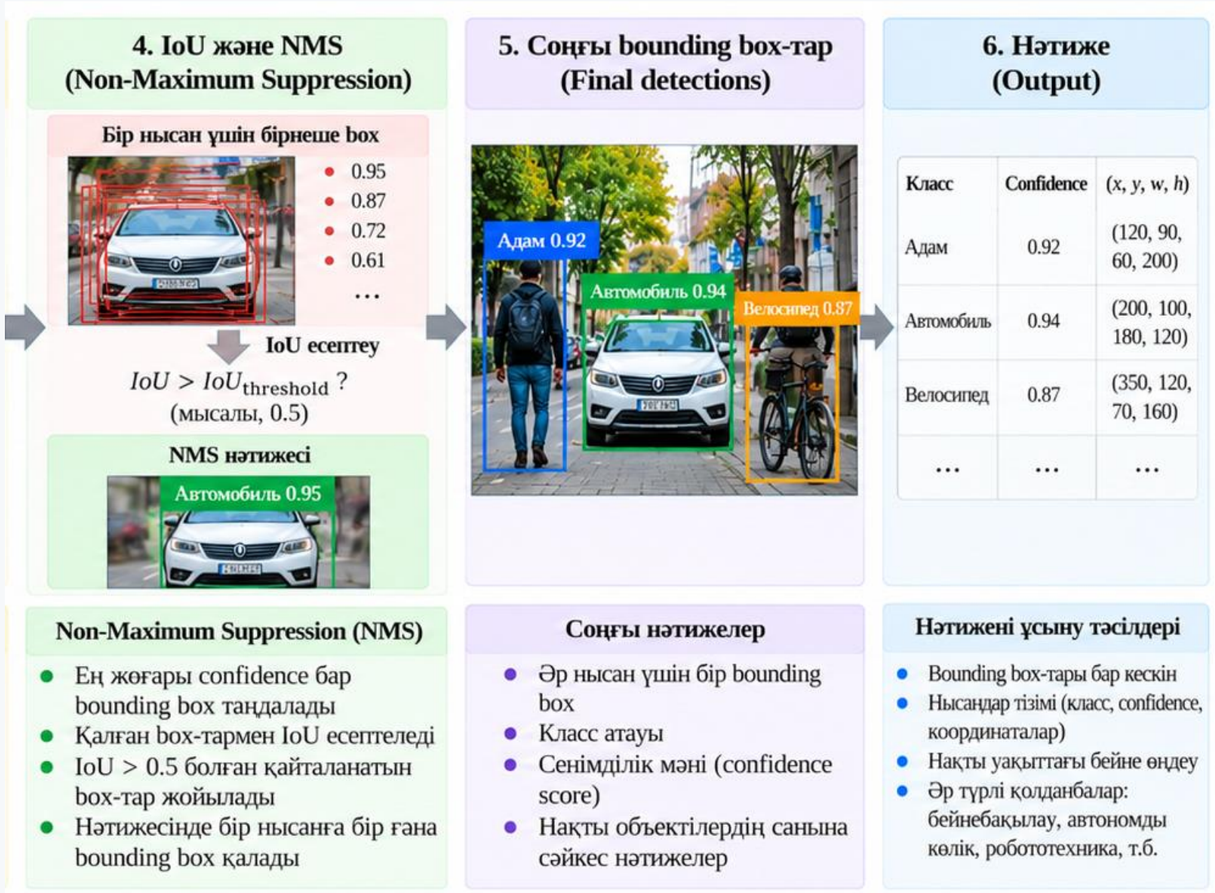
S × S grid және объект орталығына жауапты ұяшық —
классикалық түсіндірме; барлық YOLO нұсқасына бірдей ереже емес.

Нәтижені өңдеу: болжамдар және confidence сүзгісі



Дереккөз: конспект, 5-бет; 6.3-суреттің бірінші бөлігі.

Нәтижені өңдеу: NMS және соңғы анықтаулар



Суреттегі score мәндері — шартты мысалдар; NMS оларды арттырмайды.

Дереккөз: конспект, 6-бет; 6.3-суреттің екінші бөлігі.

Детекторды оқыту: аннотация және шығын

Оқыту деректері: әр кескіндегі нысандардың кластары мен нақты қораптары.

$$L = L_{box} + L_{obj} + L_{cls}$$

L_box — орналасу қатесі; L_obj — нысанның бар-жоғы;
L_cls — класс қатесі. Бұл — шығын құрамының бір үлгісі.

$$\theta \leftarrow \theta - \eta \nabla_\theta L$$

Нақты мүшелер мен олардың салмақтары YOLO нұсқасына тәуелді.

Дұрыс анықтау дегеніміз не?

TP — дұрыс анықталған нысан

Класс дұрыс.
IoU ≥ бағалау шегі.

Нақты қораппен бірмәнді сәйкестендіру жасалады.

FP және FN

FP: жалған болжам немесе қайталанатын анықтау.

FN: табылмай қалған нақты нысан.

Precision = TP / (TP + FP) Recall = TP / (TP + FN)

Average Precision және mAP

AP — бір класс үшін Precision–Recall қисығы бойынша сапа.

$$mAP = (1/C)\ \Sigma_c\ AP_c$$

mAP@0,5

Сәйкестік IoU ≥ 0,5 шегімен бағаланады.

mAP@0,5:0,95

IoU: 0,50; 0,55; …; 0,95 шектері бойынша орташа.

Модельдерді бірдей деректерде және бірдей бағалау хаттамасымен салыстырамыз.

YOLO мүмкіндіктері және практикалық шектеулері

Артықшылықтары

Бір forward pass-та көп нысанды болжау.
Көпмасштабты белгілерді пайдалану.

Бейнебақылау, өндірістік бақылау, робототехника және көлік жүйелері.

Шектеулері

Ұсақ, жабылған және тығыз орналасқан нысандар.
Жарық пен кескін сапасының өзгеруі.

Жылдамдық пен дәлдіктің ресурстарға тәуелділігі.

Бір forward pass бүкіл жүйеде қосымша өңдеу жоқ дегенді білдірмейді.

Бақылау сұрақтары

01 Classification, localization және detection есептері қалай ажыратылады?

02 Bounding box координаталары қалай беріледі және нормаланады?

03 Confidence, IoU және NMS қандай әртүрлі міндеттерді орындайды?

04 Backbone, Neck және Detection Head қызметтері қандай?

05 Детекторға қандай аннотациялар қажет және шығын нені бағалайды?

06 TP/FP/FN, AP және mAP арқылы детекция сапасы қалай бағаланады?

Архитектура: Image → Backbone → Neck → Head; өңдеу: score сүзгісі → NMS.

Дереккөз және ғылыми нақтылаулар

Негізгі дереккөз

«6-дәріс. Нысандарды анықтауға арналған нейрондық желілер. YOLO архитектурасы». Дәріс конспектісі, 9 бет.

Түпнұсқа иллюстрациялар

6.1-сурет: есептердің айырмасы, 2-бет → 2-слайд.
6.2-сурет: архитектура, 4-бет → 7-слайд.
6.3-сурет: өңдеу сатылары, 5–6-беттер → 9–10-слайдтар.

Нақтылаулар: кіріс пен feature map өлшемдерінің сәйкестігі; YOLO нұсқаларының айырмасы; NMS-тің score мәнін өзгертпеуі; дұрыс сәйкестендіру және mAP бағалау шектерін көрсету.