

YOLOV5 APMĀCĪBA ĀBOLU ATPAZĪŠANAI DĀRZĀ TRAINING OF YOLOV5 FOR APPLE DETECTION IN ORCHARD

Autori: **Aleksejs Kardjalis**, e-pasts: ak21354@edu.rta.lv,

Inga Savicka, e-pasts: is22081@edu.rta.lv,

Kaspars Rubuļņiks, e-pasts: kr08007@edu.rta.lv

Zinātniskā darba vadītājs: **Sergejs Kodors, Dr.sc.ing.**, e-pasts: sergejs.kodors@rta.lv

Rēzeknes Tehnoloģiju akadēmija, Atbrīvošanas aleja 115, Rēzekne

Abstract. Fruit cultivation is a significant part of the country economics and agriculture. In this paper, aiming to improve apple cultivation, we trained a neural network for apple detection using the YOLOv5 architecture, utilizing the dataset from the lzp-2021/1-0134 project. The dataset consisted of a set of apple tree photographs with apple fruits. Dataset contained 40 images with size 640x640 px. YOLOv5m model was trained five times. The best result model achieved mAP@0.5 equal to 0.9 and mAP@0.5:0.95 equal to 0.63. The artificial intelligence opens new possibilities for horticulture saving resources, which can be redirected on other tasks significantly increasing the efficiency of commercial orchards.

Keywords: artificial intelligence, fruits, object detection, precision horticulture, YOLOv5.

Ievads

Augļkopība ir būtiska ne tikai cilvēkiem, bet arī lopkopībā. Augļi tiek pielietoti gan cilvēka, gan mājdzīvnieku diētā. Augļkopība nes lielu vērtību ne tikai kā pārtikas avots, bet arī ekonomiskajā ziņā. Dažādi augļi tiek patērēti miljoniem cilvēku visā pasaulē, kas veicina tirgus un ekonomikas kopējo izaugsmi. Daļa ekonomikas ir balstīta uz pārtikas produktu pārdošanas, tostarp augļiem. Augļkopība un augļu pārdošana nodrošina cilvēkiem darbavietas, ienākumus un ieguldījumu pasaules ekonomikā.

Lieli komerciāli āboli dārzi vienmēr prognozē ābolu ražu, plānojot ābolu tirdzniecību. Ļoti svarīga ir precīza datu vākšana, novērtēšana un analīze. Ar moderno tehnoloģiju un mākslīgā intelekta (MI) palīdzību prognozēšanas procesus var vienkāršot. MI var apsekot lielu komerciālu augļu dārzu, saskaitīt tajā esošo koku skaitu, aprēķināt augļu skaitu konkrētā augļu dārza sektorā. Šādā veidā dārzkopji spēj optimizēt plānošanas un datu analīzes procesu un uzlabot resursu pielietojšanas efektivitāti. Ietaupītus resursus, piemēram, laiku, darbaspēku u.c., var izmantot citiem uzdevumiem, kas palielinās ābolu un citu augļu audzēšanas efektivitāti un rentabilitāti.

Neironu tīkli ir matemātiskie modeli, kas spēj imitēt cilvēka smadzeņu darbību, jo īpaši spēju mācīties uz piemēriem. MI ir vissvarīgākā tehnoloģija mūsdienu industrijā, kas spēj apstrādāt un izmantot datus dažādās dzīves nozarēs, tostarp dārzkopībā. Šajā rakstā ir aplūkota neironu tīklu apmācība ābolu skaitīšanai.

Pētījuma mērķis: apmācīt YOLOv5m modeli, kas spēj meklēt ābolus dārza fotogrāfijās.

Uzdevumi:

- 1) sagatavot datu kopu neironu tīkla apmācīšanai;
- 2) apmācīt neironu tīklu, pielietojot YOLOv5m modeli;
- 3) novērtēt apmācīta neironu tīkla precizitāti.

Materiāli un metodes

Lai apmācītu YOLOv5 neironu tīklu, tika izmantotas fotogrāfijas, kas tika savāktās lzp-2021/1-0134 projekta ietvaros. Datu kopu veido 40 640x640 pikseļu izmēra attēli, kuros attēloti koki ar āboliem. Sākumā visi attēli tika saglabāti vienā kopīgā mapē, un pēc tam, izmantojot Python skriptu, visi attēli tika nejauši sadalīti trīs dažādās mapēs. Pirmās divas mapes ir paredzētas neironu tīklu apmācībai - apmācības un validācijas kopas; un vēl viena mape ir paredzēta testēšanai. Zemāk ir parādīts piemērs ar diviem attēliem no datu kopas (sk. 1. attēlu).



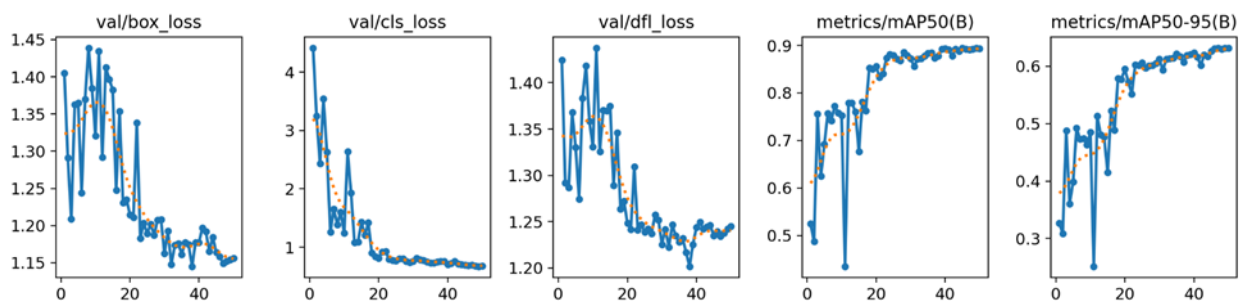
1. attēls. Attēlu piemēri no datu kopas

Datu kopa tika sadalīta trīs dažādās mapēs attiecībā 65%, 25% un 10%, atbilstoši – apmācībai, validācijai un testēšanai. Tas bija nepieciešams, lai pareizi apmācītu modeli un pārbaudītu tā atpazīšanas precizitāti. Pielietotas tehnoloģijas: dators ar NVIDIA videokarti, kas atbalsta CUDA tehnoloģiju [1], YOLOv5 satvars [2], Python un MakeSense [3].

Nākamais solis bija apmācīt neironu tīklu ar *YOLOv5m* modeli, izmantojot personīgo datoru un pielietojot *GPU (Graphics Processing Unit)* skaitļošanu. YOLOv5 neironu tīkla apmācība sastāvēja no vairākiem posmiem – YOLOv5 pakotnes tika instalētas, iekļaujot nepieciešamās atkarības, tika izvēlēts YOLOv5m modelis, YOLOv5m precizitāte tika novērtēta, tika veikta neironu tīkla apmācības progresa vizualizācija, testējot treniņus ar attēliem, rezultāti tika saglabāti secinājumu izstrādei. Lai iegūtu precīzākus treniņa rezultātus, visas šīs darbības tika veiktas 5 reizes, katru reizi mainot trenēšanas, validācijas un pārbaudes kopu attēlus.

Rezultāti un diskusija

Neironu tīklu apmācību veica piecas reizes, lai varētu vēlāk novērtēt tās efektivitāti un izvēlēties labāku rezultātu. Pēc katras apmācības mēs saņēmām grafikus ar atšķirīgiem rezultātiem. Lai turpmāk veiktu aprēķinus, mēs saglabājam testa grafikus, jo tie satur nepieciešamo informāciju, lai novērtētu cik labi tika pamācīts neironu tīkls. Zemāk ir parādīti labākie apmācību rezultāti (skat. 2. att.).



2. attēls. Apmācības rezultāti: labākais gadījums

2. attēlā ir pieci grafiki ar zaudējuma un precizitātes rezultātiem.

Zaudējuma funkcijas:

- **val/box_loss:** šī diagramma parāda modeļa veiktspēju, precīzi prognozējot robežlodziņu ap āboliem. Sākotnējā lielā mainīgums stabilizējās, kas liecina, ka, apmācībai progresējot, modelis kļūst labāks, lai attēlos lokalizētu ābolus.
- **val/cls_loss:** klasifikācijas zudumu grafiks parāda modeļa spēju pareizi klasificēt identificētos objektus. Ievērojams zaudējumu samazinājums norāda uz uzlabotu klasifikācijas precizitāti laika gaitā, norādot, ka modelis kļūst arvien prasmīgāks, lai atšķirtu ābolus no fona.
- **val/dfl_loss:** attēlojot modeļa nespēju pareizi noteikt objektus, validācijas noteikšanas

kļūmes zudums sākas augsts, taču ievērojami samazinās, ejot uz priekšu, atspoguļojot modeļa pieaugošās spējas atpazīt un noteikt ābolus dažādos apstākļos.

Precizitātes funkcijas:

- ***mAP@0.5***: vidējā precizitāte krustojumā virs Savienības (IoU) sliekšņa 0,5 saglabā augstas vērtības visā pasaulē, kas liecina, ka modelim ir spēcīga spēja pareizi noteikt ābolus, ja prognoze un pamatpatiesība pārklājas vismaz par vienu (50%).
- ***mAP@0.5:0.95***: šajā diagrammā ir sniegts precizitātes novērtējums IoU sliekšņu diapazonā no 0,5 līdz 0,95. Augšupošā tendence norāda uz modeļa palielināšanos spēju noteikt ābolus dažādos pārklāšanās līmeņos, un veikspējas līmenis samazinās, tuvojoties vēlākiem laikmetiem. Šis plato parāda, ka modelis ir sasniedzis konsekvētu noteikšanas spēju dažādās precizitātes prasību pakāpēs.

Lai saprastu, cik labi modelis atpazīst objektus, un piefiksētu precizitāti, pētījumos izmantosim *mAP@0.5:0.95* maksimālās vērtības. Zemāk var redzēt tabulu ar *mAP@0.5:0.95* maksimālām vērtībām no visām 5 apmācībām (skat. 1. tab.).

1. tabula

***mAP@0.5:0.95* maksimālās vērtības pēc katras apmācības**

Nr.	1	2	3	4	5
<i>mAP@0.5:0.95</i>	0,57579	0,57835	0,59459	0,62373	0,63259

Pēc vērtību sakārtošanas var secināt, ka:

- Minimālā *mAP@0.5:0.95* vērtība – 0,57579;
- Maksimālā *mAP@0.5:0.95* vērtība – 0,63259.

Mediāna tiek aprēķināta pielietojot nepāra skaitļa formulu (1), iegūstot rezultātā vērtību –0,59459.

$$Me = X_{\left[\frac{n+1}{2}\right]}, \quad (1)$$

kur X – sakārtots vērtību saraksts; n – vērtību skaits datu kopā.

Lai atrastu vidējo vērtību, izmanto nepieciešamo formulu (2) un iegūst rezultātu – 0,60101.

$$V = \frac{x_1 + x_2 + \dots + x_n}{n}, \quad (2)$$

kur x – vērtība no kopas X , n – vērtību skaits kopā X .

Lai pārbaudītu, cik labi strādā apmācītais neironu tīkls, tika pielietota vizuāla metode. Pielietojot apmācīto neironu tīklu, tika apstrādātas bildes. Pārskatot attēlus bija konstatēts, ka dažos attēlos neironu tīkls atzīmēja ābolus, kur droši nevar pateikt, ka tur ir ābols, kā arī ne ļoti precīzi uzzīmētas robežas. Piemēru ar šādiem attēliem var apskatīt zemāk (skat. 3. att.).



3. attēls. Ābolu atpazīšanas piemērs, pielietojot apmācīto neironu tīklu

Secinājumi

YOLOv5m neironu tīkls tika apmācīts meklēt ābolus augļu dārzu fotogrāfijās, izmantojot personālo datoru ar NVIDIA grafisko karti, kas atbalsta CUDA tehnoloģiju un izmantojot MakeSense tehnoloģiju. Eksperimentā tika izmantotas fotogrāfijas savāktas projektā Izp-2021/1-0134. Datu kopa saturēja 40 attēlus ar 640x640 pikseļu izmēru.

Šajā pētījumā katras apmācības sesijas beigās mēs iegūstam zaudējumu un precizitātes grafikus, bet mēs ņemam vērā un izmantojam maksimālo $mAP@0.5:0.95$ vērtību, lai atspoguļotu modeļa precizitāti. Galīgajā rezultātā mēs iegūvām 5 vērtības, kur minimālā vērtība ir 0,57579 un maksimālā vērtība ir 0,63259. Arī ar formulas aprēķinu palīdzību mēs uzzinājām, ka mediānas vērtība ir 0,59459, bet vidējā vērtība ir 0,60101. Pašās beigās, izmantojot iebūvēto YOLOv5 skriptu neironu tīkla testēšanai un testa attēlus, mēs veicām pēdējo testu, kurā neironu tīkls mēģināja atpazīt ābolus. Pēc testa rezultātu pārbaudes tika konstatēts, ka mākslīgais intelekts nespēja identificēt vai precīzi noteikt ābolu, ko daļēji aizsedza lapotne, vai arī ne vienmēr spēja identificēt pat acīmredzamu ābolu. Taču vairumā gadījumu neironu tīkls pareizi atpazīna un izcēla skaidras ābola robežas.

Šo modeli var uzlabot daudzos veidos. Ja ir lielāks apstrādājamo fotogrāfiju apjoms un mākslīgo intelektu var apmācīt atpazīt ābolus pat nestandarta situācijās. Šādi uzlabojumi prasīs vairāk resursu apmācībai, bet tad tas sevi attaisnos un resursu patēriņš samazināsies vairākas reizes.

Summary

Fruit production is a very important part of life, not only for humans but also for animals. Fruit is a source of sucrose, vitamins and many nutrients that are so necessary for humans. Fruit is of great value not only as a source of food but also in economic terms. Different fruits are consumed by millions of people around the world, contributing to the overall growth of the market and the economy. Part of the economy is based on the sale of food products, including fruit. Growing and selling fruit provides people with jobs, income and a contribution to the world economy. Large commercial orchards always forecast the yield of apples and other fruits, as detailed data collection, evaluation and analysis are essential when growing apples for sale.

Using neural networks to detect fruits on trees involves training the network with images

of trees and their fruits. Neural networks can be employed to identify fruit on trees by training them with images of trees and their corresponding fruits. Through computer vision techniques, the network can subsequently recognize and tally fruits by analyzing their characteristics such as shape, size, color, and texture. This paper explores the viability of utilizing the YOLOv5 neural network for apple detection in orchards.

The neural network YOLOv5 was trained on a dataset which is 40 pictures of apples on trees. In the images, we manually labeled all apples using the MakeSense image annotation tool. The training was performed on a computer with an NVIDIA graphics card supporting CUDA technology. This technique allowed the neural network to take the labeled apples in the photos as an example, thus training the artificial intelligence to recognize and label apples in videos and photos. Recognition was not successful in all cases, but average recognition accuracy was found. Using modern technologies such as YOLOv5 will advance the optimization of agricultural operations and bring great benefits in terms of saving resources, while efficiency and productivity will be increased.

In this study, at the end of each training session, we obtain visual loss and accuracy graphs, but we consider and utilize the maximum mAP@0.5:0.95 value to reflect the model's accuracy. In the final results, we obtained 5 values, where the minimum value is 0.57579 and the maximum value is 0.63259. Additionally, with the help of calculations, we found that the median value is 0.59459, but the average value is 0.60101. Ultimately, using the built-in YOLOv5 script for neural network testing and test images, we conducted the final test where the neural network attempted to identify apples on the apple tree. After verifying the test results, it was found that the artificial intelligence failed to identify or accurately determine apples partially obscured by foliage, or sometimes even failed to identify obvious apples. However, in most cases, the neural network correctly recognized and highlighted the clear boundaries of the apple.

This model can be improved in many ways. If there is a larger volume of processed photographs and artificial intelligence can be trained to recognize apples even in non-standard situations. Such improvements will require more resources for training, but then it will justify itself, and resource consumption will decrease several times.

Literatūras un avotu saraksts

1. NVIDIA. CUDA. <https://docs.nvidia.com/cuda/doc/index.html> , sk. 05.04.2024
2. Ultralytics. Train Custom Data with YoLoV5 <https://github.com/ultralytics/yolov5/wiki/Train-Custom-Data/b8cf12a92994b96f7454dc9a4f6b289f78fb9e64> , sk. 05.04.2024.
3. MakeSense. <https://www.makesense.ai>, sk. 05.04.2024.