



**Kauno technologijos universitetas**

Informatikos fakultetas

# **Automatizuotas vaisių kokybės vertinimas taikant sąsūkų neuroninius tinklus**

Baigiamasis magistro projektas

---

**Gintarė Petkutė**

Projekto autorė

**Prof. Robertas Damaševičius**

Vadovas

---

**Kaunas, 2024**



**Kauno technologijos universitetas**

Informatikos fakultetas

# **Automatizuotas vaisių kokybės vertinimas taikant sąsūkų neuroninius tinklus**

Baigiamasis magistro projektas

Dirbtinio intelekto informatika (6211BX007)

---

**Gintarė Petkutė**

**Prof. Robertas Damaševičius**

Vadovas

**Dr. Vytenis Punys**

Recenzentas

---

**Kaunas, 2024**



**Kauno technologijos universitetas**

Informatikos fakultetas

Gintarė Petkutė

## **Automatizuotas vaisių kokybės vertinimas taikant sąsūkų neuroninius tinklus**

Akademinio sąžiningumo deklaracija

Patvirtinu, kad:

1. baigiamąjį projektą parengiau savarankiškai ir sąžiningai, nepažeisdama(s) kitų asmenų autoriaus ar kitų teisių, laikydamasi(s) Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo nuostatų, Kauno technologijos universiteto (toliau – Universitetas) intelektualinės nuosavybės valdymo ir perdavimo nuostatų bei Universiteto akademinės etikos kodekse nustatytų etikos reikalavimų;
2. baigiamajame projekte visi pateikti duomenys ir tyrimų rezultatai yra teisingi ir gauti teisėtai, nei viena šio projekto dalis nėra plagijuota nuo jokių spausdintinių ar elektroninių šaltinių, visos baigiamojo projekto tekste pateiktos citatos ir nuorodos yra nurodytos literatūros sąrašė;
3. įstatymų nenumatytų piniginių sumų už baigiamąjį projektą ar jo dalis niekam nesu mokėjęs (-usi);
4. suprantu, kad išaiškėjus nesąžiningumo ar kitų asmenų teisių pažeidimo faktui, man bus taikomos akademinės nuobaudos pagal Universitete galiojančią tvarką ir būsiu pašalinta(s) iš Universiteto, o baigiamasis projektas gali būti pateiktas Akademinės etikos ir procedūrų kontrolieriaus tarnybai nagrinėjant galimą akademinės etikos pažeidimą.

Gintarė Petkutė

*Patvirtinta elektroniniu būdu*

Gintarė Petkutė. Automatizuotas vaisių kokybės vertinimas taikant sąsūkų neuroninius tinklus. Magistro baigiamasis projektas / vadovas Prof. Robertas Damaševičius; Kauno technologijos universitetas, Informatikos fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Fiziniai mokslai, Dirbtinio intelekto informatika.

Reikšminiai žodžiai: konvoliuciniai neuroniniai tinklai, kompiuterinė rega, vaisių kokybės klasifikavimas.

Kaunas, 2024. 40 p.

### **Santrauka**

Vaisių ir daržovių skynimo, sandėliavimo ir transportavimo metu gali atsirasti įvairių paviršiaus defektų, dėl ko sumažėja jų ekonominė vertė. Taip pat, neaptikus kenkėjų pažeistų vaisių ar daržovių liga ar patys kenkėjai gali būti platinami. Paprastai vaisių ir daržovių produktų defektų aptikimas ir klasifikavimas vis dar atliekamas rankiniu būdu. Tradicinis rankinis klasifikavimas pasižymi mažu tikslumu ir didelėmis sąnaudomis. Šiame darbe siekiamas efektyvus vaisių ir daržovių kokybės bei šviežumo vertinimas, naudojant giliojo mokymo metodus. Apžvelgiami konvoliuciniais neuroniniais tinklais grįsti metodai, jie tarpusavyje palyginami bei pasiūloma architektūros modifikacija.

Gintarė Petkutė. Automated fruit quality evaluation using convolutional neural networks. Master's Final Degree Project / supervisor Prof. Robertas Damaševičius; Faculty of Informatics, Kaunas University of Technology.

Study field and area (study field group): Physical science, Artificial Intelligence in Computer Science.

Keywords: convolutional neural networks, computer vision, fruit quality evaluation.

Kaunas, 2024. 40 p.

### **Summary**

Fruit and vegetables can suffer from various surface defects during picking, storage and transport, which can reduce their economic value. In addition, the absence of detection of pest damage to the fruit or vegetables may lead to the spread of the disease or the pests themselves. Generally, the detection and classification of defects in fruit and vegetable products is still done manually. Traditional manual grading is characterised by low accuracy and high cost. This work aims at efficient assessment of quality and freshness of fruits and vegetables using deep learning techniques. Convolutional neural network based methods are reviewed, compared with each other and an architectural modification is proposed.

## Turinys

Lentelių sąrašas .....	7
Paveikslų sąrašas .....	8
Santrumpų ir terminų sąrašas .....	9
Įvadas.....	10
<b>1. Vaisių kokybės vertinimo uždavinio analizė.....</b>	<b>11</b>
1.1. Vaisių atpažinimo ir kokybės vertinimo būdai.....	11
1.2. Neuroniniai tinklai.....	11
1.3. Aktyvacijos funkcijos .....	12
1.4. Konvoliuciniai neuroniniai tinklai.....	13
1.4.1. Konvoliucijos sluoksnis.....	14
1.4.2. Maksimalus telkimas .....	14
1.4.3. Išplėtimas.....	15
1.5. Giliuoju mokymu grįsti objektų aptikimo metodai .....	15
1.6. Giliuoju mokymu grįsti klasifikavimo metodai .....	16
1.6.1. „AlexNet“ .....	16
1.6.2. „Vision Transformers“ .....	17
1.6.3. „MobileNetV3“ .....	18
1.6.4. „EfficientNetV2“ .....	19
1.7. Modelių palyginimas .....	19
1.8. Atviros prieigos duomenų rinkiniai, skirti vaisių kokybės įvertinimui .....	20
<b>2. Automatizuoto vaisių kokybės tyrimo projektavimas .....</b>	<b>23</b>
2.1. Sistemos koncepcija .....	23
2.2. Funciniai reikalavimai .....	24
2.3. Nefunkciniai reikalavimai .....	24
2.4. Aparatūrinė įranga .....	24
2.5. Sprendimo kūrimo priemonės .....	25
2.6. Vertinimo metrikos.....	25
2.6.1. Tikslumas, preciziškumas, jautrumas.....	25
2.6.2. mAP50 įvertis.....	26
<b>3. Eksperimentinė dalis .....</b>	<b>27</b>
3.1. Duomenų paruošimas .....	27
3.1.1. Duomenų paruošimas objektų aptikimo uždaviniui .....	27
3.1.2. Duomenų paruošimas kokybės klasifikavimo uždaviniui.....	27
3.1.3. Duomenų paruošimas modelių rezultatų testavimui su nematytais duomenimis.....	28
3.2. Atrinktų modelių mokymas .....	30
3.3. Modelių rezultatų analizė .....	32
3.4. Modelių pritaikomumo naujiems duomenų rinkiniams įvertinimas .....	35
3.4.1. Vaisių aptikimo ir rūšies klasifikavimo tikslumas .....	35
3.4.2. Vaisių kokybės vertinimo tikslumas .....	36
<b>Išvados .....</b>	<b>38</b>
<b>Literatūros sąrašas .....</b>	<b>39</b>

## Lentelių sąrašas

<b>1 lentelė.</b> Modelių parametrų skaičiaus palyginimas .....	19
<b>2 lentelė.</b> Vaisių duomenų rinkinių palyginimas .....	20
<b>3 lentelė.</b> Naudojamų augmentacijų parametrai .....	27
<b>4 lentelė.</b> „YOLOv8“ skirtingų modelio versijų hiperparametrai .....	30
<b>5 lentelė.</b> Kokybės klasifikavimo modelių hiperparametrai .....	31
<b>6 lentelė.</b> Transformerių modelio hiperparametrai .....	32
<b>7 lentelė.</b> „YOLOv8“ skirtingų modelių mokymo rezultatai .....	32
<b>8 lentelė.</b> Kokybės klasifikavimo modelių rezultatai .....	33
<b>9 lentelė.</b> Vaisių šviežumo vertinimo sistemos tikslumas su nematytais duomenimis .....	36

## Paveikslų sąrašas

<b>1 pav.</b> Paprastas neuroninis tinklas .....	12
<b>2 pav.</b> ReLU funkcija [8] .....	12
<b>3 pav.</b> Sigmoid funkcija [8].....	13
<b>4 pav.</b> Konvoliucinio neuroninio tinklo sandara .....	14
<b>5 pav.</b> Konvoliucijos operacija.....	14
<b>6 pav.</b> Maksimalaus telkimo operacija.....	15
<b>7 pav.</b> Transponuotos konvoliucijos operacija.....	15
<b>8 pav.</b> Palyginimas tarp „YOLOv8“ ir kitų YOLO modelių [15].....	16
<b>9 pav.</b> „AlexNet“ architektūra [17] .....	17
<b>10 pav.</b> Vaizdo transformeriai [18] .....	18
<b>11 pav.</b> „MobileNetV3“ architektūros elementas [20].....	18
<b>12 pav.</b> „EfficientNetV2“ architektūra [21] .....	19
<b>13 pav.</b> „FruitQ“ duomenų rinkinio atsitiktinai parinktos nuotraukos.....	21
<b>14 pav.</b> „FruitQ“ duomenų rinkinio klasių pasiskirstymas .....	22
<b>15 pav.</b> Automatizuoto vaisių kokybės vertinimo tyrimo panaudos atvejų diagrama .....	23
<b>16 pav.</b> Detalizuota sistemos koncepcija .....	24
<b>17 pav.</b> Sumaišymo matricos pavyzdys .....	25
<b>18 pav.</b> „CVAT“ programos duomenų anotavimo langas .....	27
<b>19 pav.</b> Pritaikytų augmentacijų pavyzdžiai .....	28
<b>20 pav.</b> Naujai sukurto duomenų rinkinio visų klasių atsitiktinai parinktos nuotraukos .....	29
<b>21 pav.</b> Naujai sukurto duomenų rinkinio pasiskirstymas .....	30
<b>22 pav.</b> Geriausio „YOLOv8“ modelio rezultatai .....	33
<b>23 pav.</b> „EfficientNetV2B0“ modelio rezultatų reagavimo zonos.....	34
<b>24 pav.</b> v4.0 versijos objektų aptikimo modelio klaidų matrica su nematytais duomenimis.....	35
<b>25 pav.</b> Vaisių šviežumo vertinimo sistemos tikslumas su nematytais duomenimis.....	37

## Santrumpų ir terminų sąrašas

### Santrumpos:

ReLU (angl. *Rectified Linear Unit*) – tiesinė aktyvacijos funkcija.

CNN (angl. *Convolutional Neural Network*) – konvoliucinis neuroninis tinklas.

mAP (angl. *Mean Average Precision*) – vidutinio tikslumo vidurkis.

IoU (angl. *Intersection Over Union*) – aibių sankirtos bei sąjungos santykio vertė.

YOLO (angl. *You Only Look Once*) – gilioju mokymu grįstas objektų aptikimo, realiuoju laiku, metodas.

## Įvadas

Žemės ūkio pramonė atlieka svarbų vaidmenį daugelio šalių ekonomikoje. Svarbus žemės ūkio sektorius yra šviežių vaisių ir daržovių gamyba ir tiekimas pardavėjams. Didėjanti efektyvios maisto produktų gamybos ir greito bei saugaus tiekimo paklausa, paskatino šioje pramonėje naudoti įvairias naujoviškas technologijas [1]. Tikslas yra sukurti intelektualias vaisių ir daržovių atrinkimo sistemas, kurios būtų pigios ir lengvos naudotis.

Taip pat, didėjant populiacijai, didėja ir maisto švaistymas. Kasmet pasaulyje yra išmetama 1,3 milijardai tonų dar tinkamo vartoti maisto atliekų. Šios maisto atliekos į aplinką išmeta 3,3 milijonų tonų CO<sub>2</sub> [2], o tai prilygsta į atmosferą išmetamų šiltnamio efektą sukeliančių dujų kiekiui. Dažnai į šaldytuvą įdėtos daržovės ar vaisiai būna pamiršti ir jie sugenda. Vienas iš vaisių ir daržovių kokybės atpažinimo pritaikymų leistų išmaniesiems šaldytuvams atsiųsti pranešimą į telefoną ar išmanųjį laikrodį apie senti pradedančius produktus [3]. Tačiau, vaisių ir daržovių atpažinimo užduotis nėra lengva dėl didelės įvairovės skirtingų formų, spalvų, dydžių, tekstūros ir kitų vaisių bei daržovių charakteristikų.

Šiame darbe nagrinėsime kaip, panaudojant dirbtinį intelektą, galėtume atpažinti vaisius ir daržoves bei įvertinti jų kokybę ir šviežumą.

**Pagrindinis tikslas** – išanalizuoti ir palyginti giliojo mokymo metodus, pagrįstus konvoliuciniais neuroniniais tinklais, siekiant efektyvaus vaisių kokybės ir šviežumo įvertinimo.

### **Iškelti uždaviniai tikslo pasiekimui:**

1. išnagrinėti įvairius giliuosius neuroninius tinklus ir jų taikymą vaisių defektų aptikimui;
2. palyginti skirtingų architektūrų rezultatus ir pasirinkti tinkamiausią šios problemos sprendimui;
3. atlikti modifikacijas architektūroje, kurios leistų išgauti tikslesnius rezultatus vaisių šviežumo atpažinimui;
4. ištestuoti sistemos tikslumą su naujais modeliui nematytais duomenimis.

## 1. Vaisių kokybės vertinimo uždavinio analizė

Šiame skyriuje atliekama literatūros apžvalga, kurios metu susipažįstama su vaisių kokybės vertinimo sprendimais, konvoliuciniais neuroniniais tinklais bei jų atliekamomis pagrindinėmis operacijomis. Pristatomos kelios giliųjų neuroninių tinklų architektūros, pritaikomos šviežių ir defektuotų vaisių klasifikavimui ir aptikimui. Taip pat, apžvelgiami egzistuojantys duomenų rinkiniai naudojami šios problemos sprendimui.

### 1.1. Vaisių atpažinimo ir kokybės vertinimo būdai

Vaisių aptikimas, identifikavimas bei jų kokybės įvertinimas yra labai svarbūs etapai maisto pramonėje. Tradiciškai šie procesai atliekami rankiniu būdu, tačiau tai užima daug laiko. Monotoniškai atliekant tokį darbą neapsieinama ir be žmogiškųjų klaidų. [4] Siekiant paspartinti vaisių rūšiavimą, sumažinti darbo kaštus ir sumažinti atrinkimo klaidų tikimybę, kuriami įvairūs automatizuoti rūšiavimo metodai.

Šiuos metodus galima suskirstyti į nevizualinius ir vizualinius metodus. Taikant nevizualinius metodus norint nustatyti vaisiaus kokybę stebimos tokios savybės, kaip kvapas, cheminė sudėtis, jausmas liečiant tiriamą objektą. Kita vertus, vizualiniai metodai remiasi skaitmeniniais vaizdais, iš kurių galima ištraukti tokias savybes, kaip spalva, dydis, forma. Pastaraisiais metais šiuo metodu atliekamos įvairios analizės skirtos įvertinti vaisių kokybę [5].

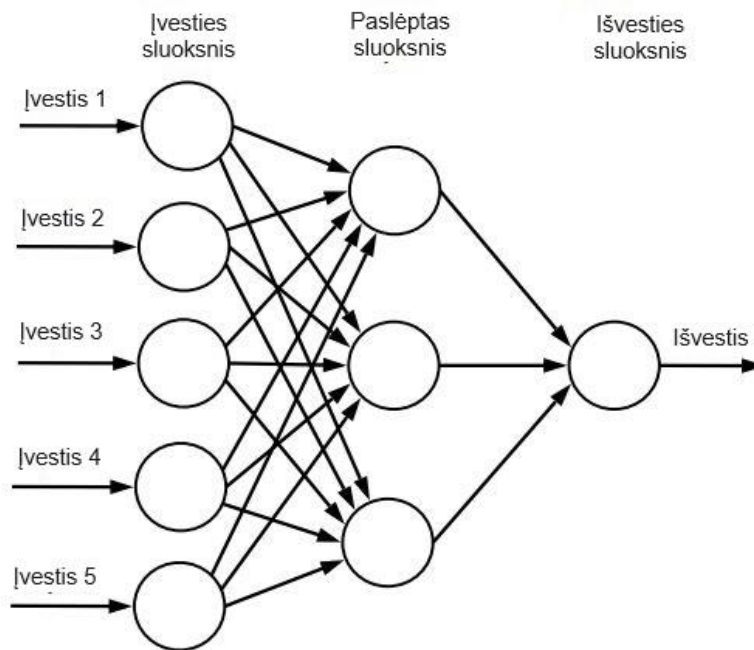
Pasitelkiant vizualines vaisių savybes jiems aptikti ir klasifikuoti taikomi įvairūs mašininio mokymosi metodai, įskaitant atraminių vektorių mašinas (angl. *Support Vector Machine*, SVM), k artimiausių kaimynų (angl. *K-Nearest Neighbours*, KNN) ir konvoliucinius neuroninius tinklus (angl. *Convolutional Neural Network*, CNN). Iš šių metodų CNN pasiekia didžiausią klasifikavimo tikslumą [6].

### 1.2. Neuroniniai tinklai

Dirbtinis neuroninis tinklas yra matematinis modelis, sukurtas pagal žmogaus biologinės nervų sistemos veikimo principus. Jį sudaro daugybė tarpusavyje susijusių dirbtinių neuronų, kurie mokosi iš įvesties duomenų ir stengiasi optimizuoti išvesties rezultatą [7].

Neuroninis tinklas susideda iš trijų pagrindinių dalių: įvesties sluoksnio, paslėptų sluoksnių ir išvesties sluoksnio. Įvesties duomenys dauginami iš pradinių svorių (angl. *weights*), o gautos sandaugos susumuojamos. Ši suma perduodama į aktyvacijos funkciją, kurią galima pasirinkti priklausomai nuo sprendžiamos problemos.

Procesas kartojamas per visus paslėptus sluoksnius, kol pasiekiamas išvesties sluoksnis. Čia gali būti naudojama dar viena aktyvacijos funkcija, kuri gali skirtis nuo ankstesnės. Galiausiai, gautos išvesties reikšmės lyginamos su faktinėmis reikšmėmis ir apskaičiuojamas klaidos įvertis. Remiantis šiuo klaidos įverčiu, svoriai koreguojami ir procesas kartojamas, siekiant pagerinti tinklo tikslumą.

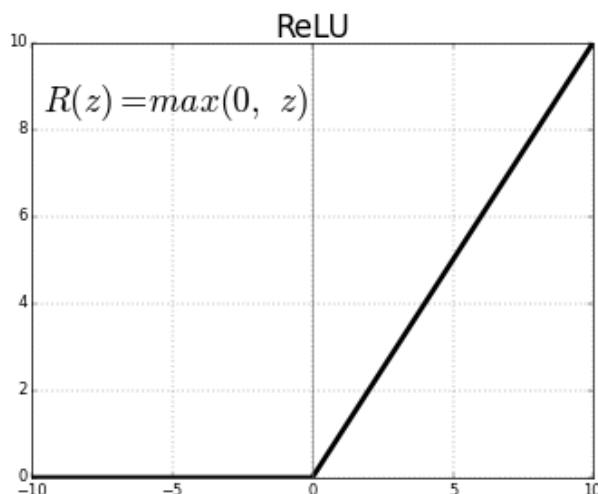


1 pav. Paprastas neuroninis tinklas

### 1.3. Aktyvacijos funkcijos

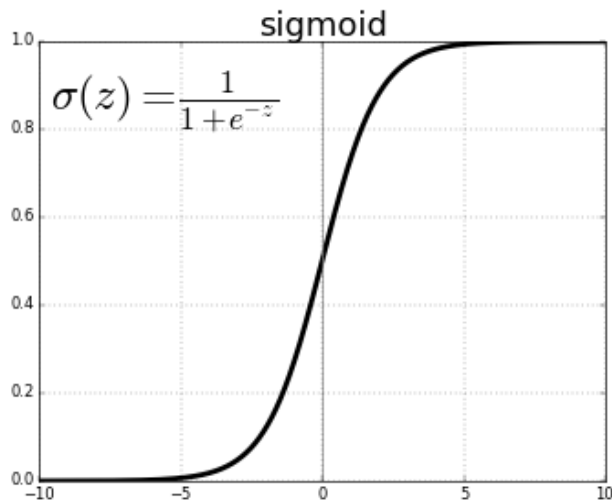
Aktyvacijos funkcijos suteikia neuroniniam tinklui gebėjimą mokytis ir interpretuoti sudėtingą duomenų informaciją. Jų pagrindinis tikslas yra įdiegti netiesiškumo savybę į tinklo veikimą.

ReLU funkcija yra viena iš tokių funkcijų, kuri ne tik suteikia netiesiškumo savybę, bet ir padeda spręsti nykstančio gradiento problemą [8]. Ji veikia taip: jei įvesties reikšmė yra neigiama, grąžinama 0, o jei reikšmė teigiama, grąžinama ta pati reikšmė. ReLU išvesties reikšmės yra tarp 0 ir begalybės.



2 pav. ReLU funkcija [8]

Sigmoid funkcija naudojama konvoliucinio neuroninio tinklo pabaigoje. Ši funkcija naudojama binarinio klasifikavimo uždaviniams, išvesties rezultatas visada bus tarp 0 ir 1 [8]. Jei reikšmė yra mažesnė nei 0,5 tuomet skaitoma, jog tai yra „0“ klasė, jei reikšmė didesnė nei 0,5, tuomet „1“ klasė.



3 pav. Sigmoid funkcija [8]

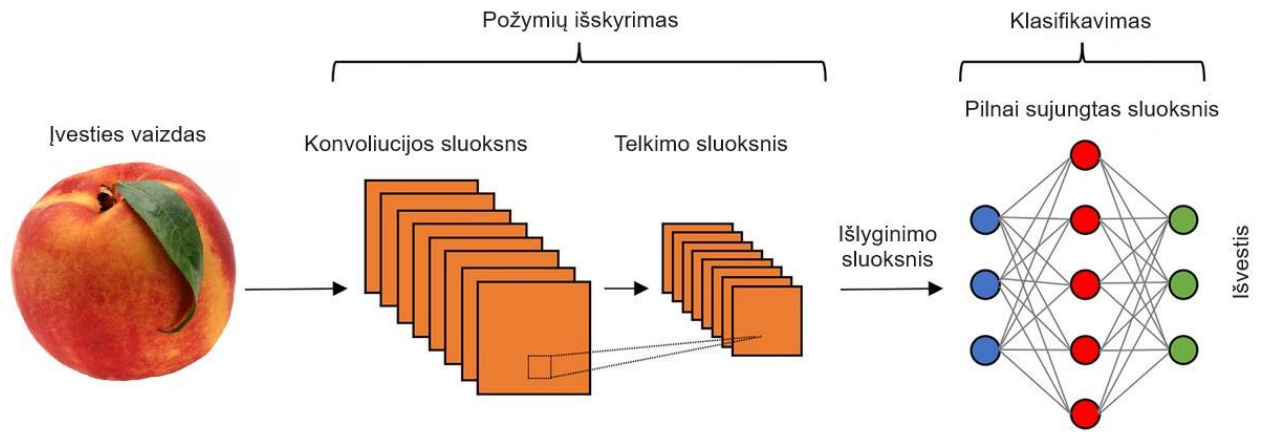
#### 1.4. Konvoliuciniai neuroniniai tinklai

Konvoliuciniai neuroniniai tinklai yra prižiūrimo giliojo mokymosi metodas, dažniausiai taikomas vaizdų atpažinimo ir kompiuterinės regos srityse. Jie yra sudėtingesni nei tradiciniai dirbtiniai neuroniniai tinklai, nes turi daugiau paslėptų sluoksnių, nors mokymosi principas išlieka panašus.

Konvoliucinių neuroninių tinklų struktūrą sudaro keli pagrindiniai komponentai [9]:

- **Įvesties sluoksnis:** gauna pradinis duomenis (pvz., vaizdą).
- **Konvoliucijos sluoksniai:** šie sluoksniai taiko filtrus vaizdo regionams, siekiant išgauti svarbias savybes ir bruožus.
- **ReLU aktyvavimo funkcijos:** kiekviename konvoliuciniame sluoksnyje naudojamos ReLU funkcijos, kurios padeda tinklui mokytis netiesiškumo.
- **Telkimo sluoksniai** (angl. *pooling*): sumažina matricos dydį, išlaikant svarbiausias savybes, taip mažinant skaičiavimo sudėtingumą ir padidinant modelio efektyvumą.
- **Visiškai sujungtas sluoksnis** (angl. *fully-connected*): jungia visus neuronus iš ankstesnio sluoksnio su visais neuronais kitame sluoksnyje, leidžiant tinklui priimti galutinius sprendimus.
- **Išvesties sluoksnis:** pateikia galutinius klasifikacijos rezultatus arba kitus sprendimus.

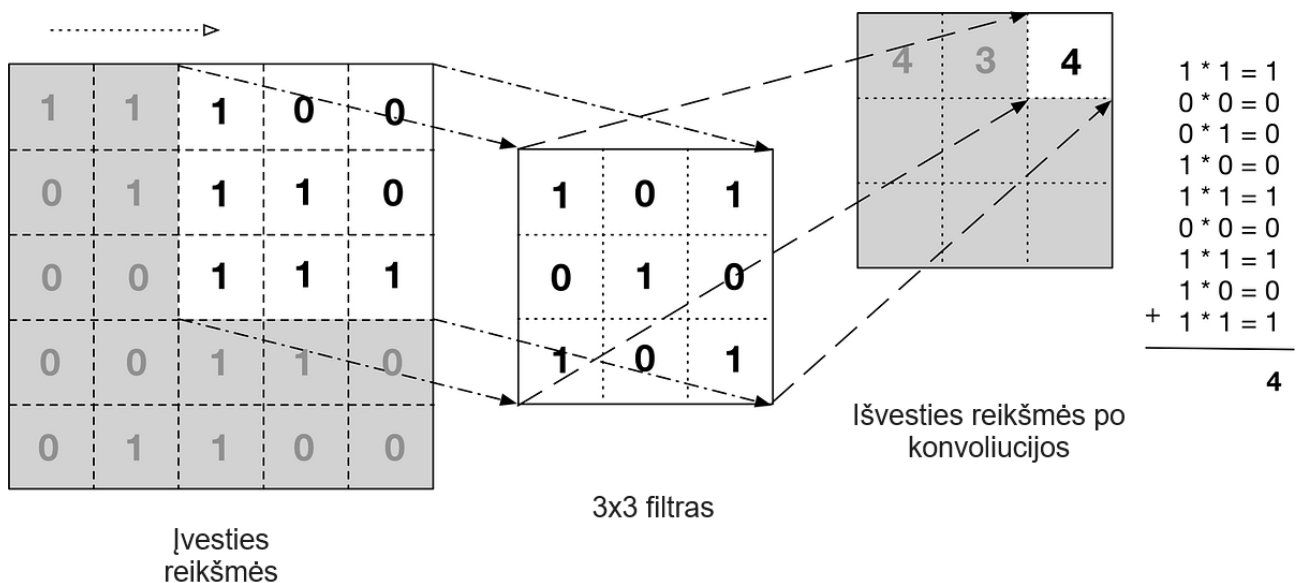
Šie komponentai kartu sudaro neuroninio tinklo struktūrą, kuri geba efektyviai analizuoti ir atpažinti vaizdus.



4 pav. Konvoliucinio neuroninio tinklo sandara

### 1.4.1. Konvoliucijos sluoksnis

Konvoliucijos sluoksniai turi nustatytą pradinį filtrų skaičių, kuris kartu su jų svoriais kinta mokymo metu. Filtrų dydžiai dažniausiai būna 3x3 arba 5x5, o jų gylis atitinka ankstesnio sluoksnio gylį. Šie filtrai juda per vaizdą iš viršaus kairės į dešinę ir žemyn, atlikdami konvoliucijos operaciją kiekviename žingsnyje. Detaliau konvoliucijos operacija parodyta 5 pav. Žingsnio (angl. *stride*) dydis taip pat yra pasirenkamas. Kai filtrai pereina visą vaizdą, kiekvienam filtrui sugeneruojamas požymių žemėlapis (angl. *feature map*). Požymių žemėlapiai perduodami aktyvacijos funkcijai, kuri nustato, ar tam tikroje vaizdo vietoje yra atitinkamas požymis [10][11].



5 pav. Konvoliucijos operacija

### 1.4.2. Maksimalus telkimas

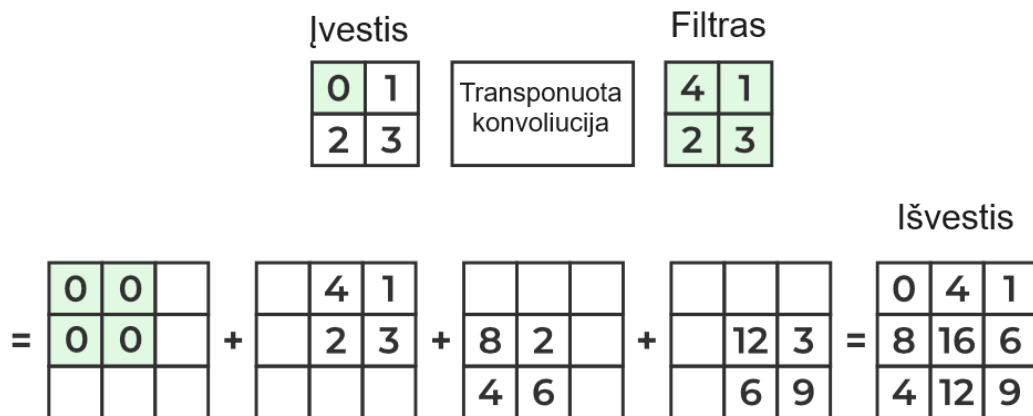
Maksimalaus telkimo (angl. *max pooling*) sluoksnis taikomas po konvoliucijos sluoksnio, kad sumažintų įvesties matricos dydį ir sumažintų apskaičiuojamų parametru kiekį. Ši operacija yra paprasta – iš matricos reikšmių, patenkančių į maksimalaus telkimo filtro langą, išrenkama didžiausia reikšmė ir perkeliama į naują duomenų matricą [12].



6 pav. Maksimalaus telkimo operacija

### 1.4.3. Išplėtimas

Išplėtimo (angl. *upsampling*) operacijos tikslas yra atkurti sumažinto vaizdo dydį į pradinį dydį. Yra keletas būdų šiai operacijai atlikti, vienas iš jų – transponuota konvoliucija (angl. *transposed convolution*). Šio metodo metu vaizdo, kurį norime padidinti, matricos kiekviena reikšmė dauginama su filtro reikšmėmis, o persidengiančios reikšmės galutiniame rezultate susumuojamos [13].



7 pav. Transponuotos konvoliucijos operacija

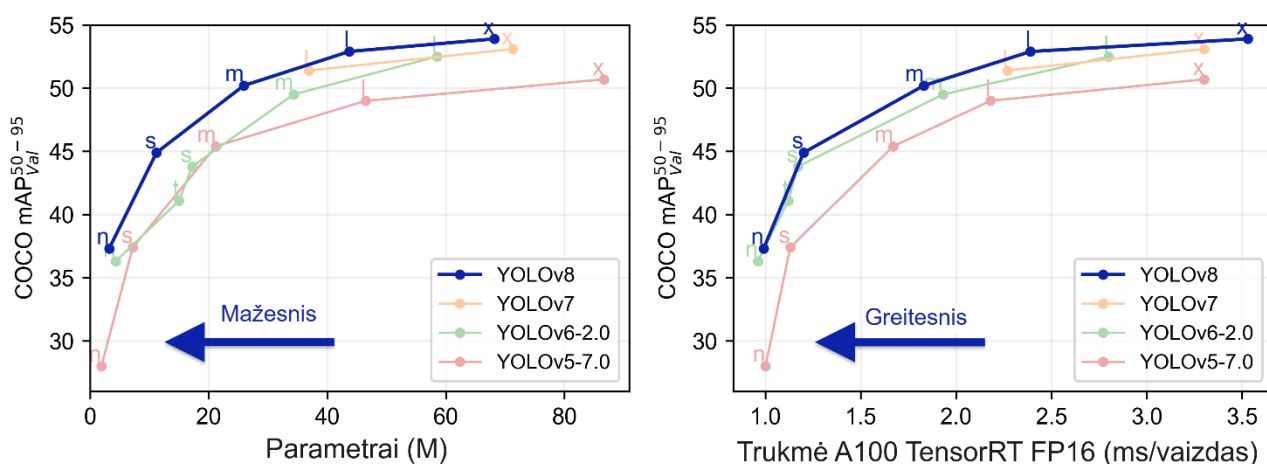
### 1.5. Giliuoju mokymu grįsti objektų aptikimo metodai

Objektų aptikimas yra pirmas žingsnis sprendžiant šią problemą, kadangi vaizde ar video medžiagoje vaisių gali būti keletas, todėl aktualu juos lokalizuoti individualiai ir nustatyti, kas tai per vaisius. Sukurtas keletas, giliuoju mokymu paremtų, objektų aptikimo metodų. Vienas iš tokių metodų pasirinktas *You Only Look Once* (YOLO) [15] dėl savo našumo, efektyvumo ir gebėjimo veikti realiu metu.

Pagrindinė YOLO metodo idėja slypi šio metodo pavadinime, jis paduodamą vaizdą apdoroja tik vieną kartą, tuom skiriasi nuo tradicinių objektų aptikimo metodų. Šis metodas vaizdus perduoda vieną kartą per pilną konvoliucinį neuroninį tinklą (angl. *Fully Convolutional Network FCN*) [15] ir

gražina aptikto objekto koordinates ir klasifikavimo prognozes vienu perdavimu, todėl šis tinklas veikia sparčiai.

Greitis yra vienas iš pagrindinių YOLO privalumų, todėl jis tinkamas taikomosioms programoms, kuriose reikia greitai ir tiksliai aptikti objektus, pavyzdžiui, greitai besikeičiančių sąlygų vaizdo stebėjimas, autonominių automobilių ir papildytosios realybės programoms. Šis metodas gali apdoroti iki 45 kadrų per sekundę, todėl puikiai tinkamas naudoti realiuoju laiku. Kitas privalumas yra tas, kad metodas ne tik greitas, bet sugeba pasiekti didesnę tikslumą nei kiti šiuolaikiniai metodai. Naujusias YOLOv8 modelis aplenkia ankstesnes šio metodo versijas, palyginimą galime matyti žemiau esančiame, 8 pav.



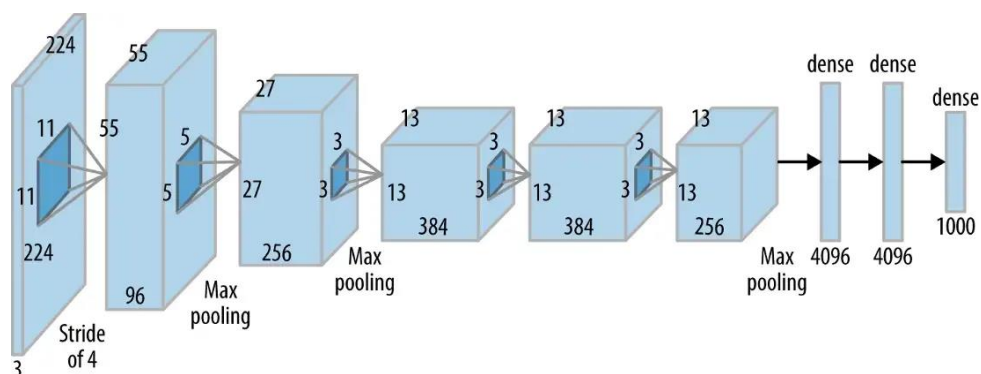
8 pav. Palyginimas tarp „YOLOv8“ ir kitų YOLO modelių [15]

## 1.6. Giliuoju mokymu grįsti klasifikavimo metodai

### 1.6.1. „AlexNet“

„AlexNet“ laikomas vienu iš svarbiausių modelių, padariusių didelę įtaką kompiuterinei regai. 2012 metais šis modelis laimėjo prestižinį kompiuterinės regos konkursą „ImageNet“, pasiekdamas 84,7% tikslumą.

Modelio mokymo sparta buvo gerokai didesnė dėl dviejų priežasčių: 1) vietoj įprastinių sigmoidinės ar hiperbolinio tangento funkcijų buvo naudojama ReLU aktyvacijos funkcija; 2) konvoliucijos metu buvo ieškomas lokalus maksimumas, o ne lokalus vidurkis [16].



9 pav. „AlexNet“ architektūra [17]

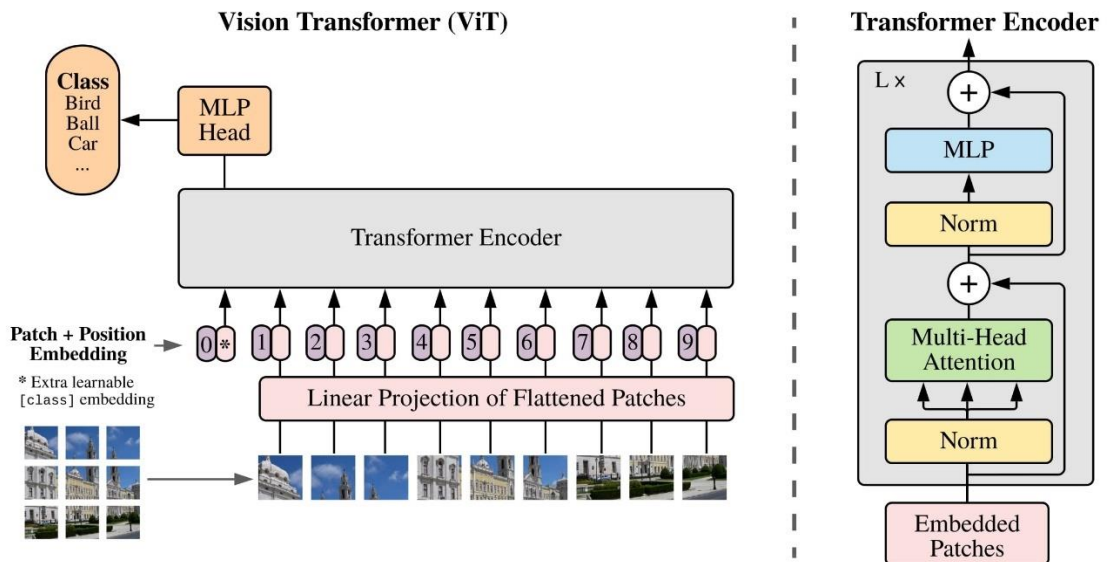
„AlexNet“ turi penkis skirtingų matmenų neuroninių tinklų sluoksnius, kuriuose tarp sujungtų sluoksnių naudojamos ReLU aktyvacijos funkcijos. Modelio įvestis yra RGB vaizdas su 227x227 raiška.

Pirmajame sluoksnyje atliekamos 96-ios 11x11 dydžio filtrų konvoliucijos per visą įvesties vaizdą. Konvoliucijos operacija atliekama „slenkančio lango“ principu, filtrą paslenkant 4 pikseliais. Rezultatai perduodami į antrą sluoksnį su 55x55x96 matmenimis. Kituose sluoksniuose procesas kartojamas su skirtingais matmenimis.

### 1.6.2. „Vision Transformers“

„Google Research“ tyrėjai *Alexey D. et al.* 2020 metais pristatė transformeriais grįstą architektūrą vaizdų aptikimui pavadinimu „Vision Transformers“ [19].

„Vision Transformers“ naudoja transformerių architektūrą, kuri originaliai sukurta natūralios kalbos užduotims, tačiau ją pritaikė vaizduo duomenims apdoroti ir analizuoti. Pirmuoju žingsniu, originalaus dydžio įvesties vaizdas yra padalijamas į fiksuoto dydžio kvadratėlius (pvz., 16x16 pikselių). Kvadratėliai laikomi „žodžiu“, analogiškai teksto duomenų ženklams. Toliau kiekvienas padalytas kvadratėlis yra suplokštinamas į 1D erdvę ir tada tiesiškai įterpiamas į fiksuoto dydžio įterpimo (angl. *embedding*) vektorių. Šie įterpimo vektoriai yra perduodami į standartinį transformerių enkoderį, kurį sudaro keletas sluoksnių kelių galvų (angl. *multi-head*), savarankiško dėmesio (angl. *self-attention*) ir tiesioginio perdavimo (angl. *feed-forward*) neuroninių tinklų.

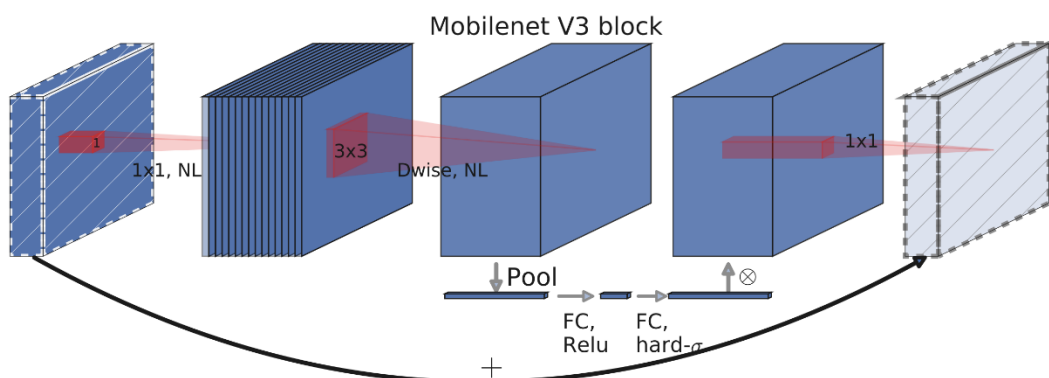


10 pav. Vaizdo transformeriai [18]

„Vision Transformers“ buvo naudojami vaisių kokybės vertinimo uždaviniams spręsti. *Bingjie Xiao et al.* 2023 metais išpublikavo straipsnį [24], kuriame vertinimo vaisių prinokimą ir jų pritaikytas vaizdų transformeris pasiekė 87,43% tikslumą obuolių ir kriaušių klasifikavime.

### 1.6.3. „MobileNetV3“

Bendrosios paskirties kompiuterinės regos neuroniniai tinklai, sukurti mobiliems įrenginiams, skirti klasifikavimui, aptikimui ir kt. Galimybė paleisti giliuosius tinklus asmeniniuose mobiliuosiuose įrenginiuose pagerina naudotojų patirtį, suteikia galimybę jais naudotis bet kuriuo metu ir bet kur, o tai suteikia papildomos naudos saugumo, privatumo ir energijos suvartojimo srityse. Atsirandant naujoms taikomosioms programoms, leidžiančioms naudotojams sąveikauti su realiuoju pasauliu realiu laiku, didėja ir vis efektyvesnių neuroninių tinklų poreikis.

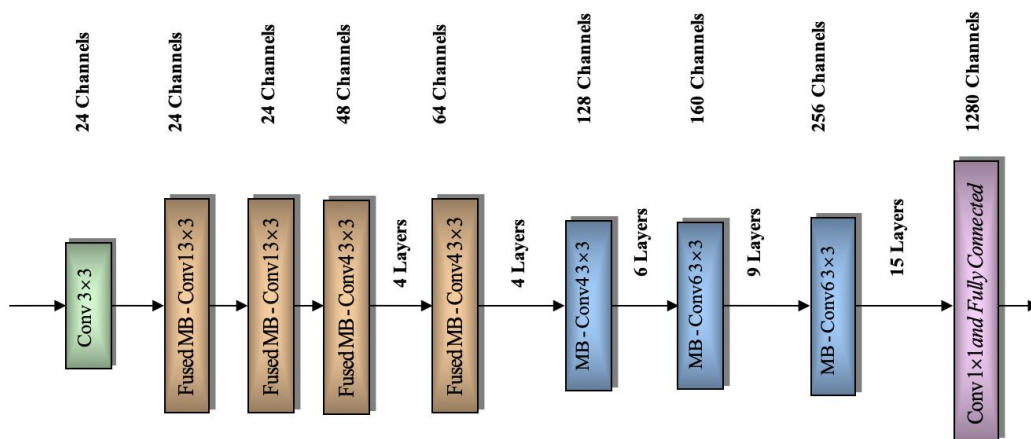


11 pav. „MobileNetV3“ architektūros elementas [20]

„MobileNetV3“ yra „Google Research“ sukurtos „MobileNet“ modelių serijos, skirtos neuroninių tinklų optimizavimui mobiliems įrenginiams, dalis. Ši modelio versija buvo pristatyta *Andrew H. et al.* [20] 2019 metais. Vienas iš pokyčių nuo buvusių šio modelio versijų, tai jįg vietoj tradicinės ReLU aktyvacijos funkcijos naudojama efektyvesnė H-Swish funkcija, kuri pagerina modelio veikimą ir sumažina skaičiavimo sąnaudas.

### 1.6.4. „EfficientNetV2“

Mingxing T. ir Quoc V. Le. [21] 2021 metais pristatė naujesnę „EfficientNet“ architektūros versiją – „EfficientNetV2“. Ši naujesnė versija sukurta siekiant pagerinti ankstesnės versijos trūkumus, daugiausia dėmesio skiriant spartesniam modelio mokymo laikui ir geresniam rezultatui išgauti naudojant mažesnę kiekį parametrų. V2 versijoje patobulintas mokymo metodas, kuomet naudojamas laipsniškas mokymasis, kai mokymo metu palaipsniui didinamas vaizdo dydis ir modelio talpa. Toks metodas leidžiai modeliui greičiau konverguoti ir padidinti tikslumą. Modelis, taip pat, taiko įvairius regulizavimo metodus, pavyzdžiui, stochastinio gylio ir duomenų augmentavimo technikas su tikslu išvengti per didelio modelio persimokymo (angl. *overfitting*).



12 pav. „EfficientNetV2“ architektūra [21]

### 1.7. Modelių palyginimas

Šiame projekte siekiama, jog pasirinktas modelis atliktų kuo tikslesnius skaičiavimus sunaudodamas kuo mažiau resursų. Iš visų aprašytų modelių didžiausią parametrų kiekį turi „Vision Transformer“ architektūros modelis.

Tolesniems eksperimentams atlikti pasirinktos 1 lentelėje išvardytos architektūros. Atmesta tik „AlexNet“ architektūra dėl didelio parametrų kiekio, kuomet literatūros apžvalgoje mažiau parametrų turinčios architektūros pasiekdavo geresnius rezultatus. „Vision Transformers“ nors ir turinti didesnę parametrų skaičių nei „AlexNet“ įtraukta į eksperimentinę dalį. Tikslas palyginti ir išsiaiškinti ar vaisių kokybės vertinimo uždaviniui spręsti vaizdų transformeriai turi pranašumą prieš mažesnius tinklus, tokius kaip „EfficientNetV2B0“ ir „MobileNetV3“.

1 lentelė. Modelių parametrų skaičiaus palyginimas

Architektūra	Parametrų skaičius
„Vision Transformers“	85,013,312
„AlexNet“	61,100,344

„EfficientNetV2B3”	15,603,688
„EfficientNetV2B0”	4,151,464
„MobileNetV3”	3,919,472

### 1.8. Atviros prieigos duomenų rinkiniai, skirti vaisių kokybės įvertinimui

Pastaruoju metu įvairioms užduotims, įskaitant vaisių aptikimą ir kokybės vertinimą, buvo pristatyti keli vaisių duomenų rinkiniai.

*Mihai O.* [22] 2021 metais pristatė duomenų rinkinį pavadinimu *Fruits360*. Duomenų rinkinį sudaro gausi įvairovė vaisių bei riešutų nuotraukų. Iš viso duomenų rinkinys turi 131 klasę, bendrai jį sudaro 90380 nuotraukų, kurių dydis yra 100x100 pikselių. Kiekvienas vaisius suskirstytas į atskirus katalogus pagal pavadinimą, todėl šie duomenys tinkami vaisių rūšies klasifikavimo uždaviniams spręsti.

*Vishal M. ir Kailas P.* [23] 2022 metais pristatė indiškų vaisių duomenų rinkinį skirtą klasifikavimo bei kokybės vertinimo uždaviniams spręsti. Šį duomenų rinkinį sudaro šešių rūšių vaisiai: obuolys, bananas, gvajava, laimas, apelsinas, granatas. Šie vaisiai suskirstyti į tris katalogus: geros kokybės vaisiai, blogos kokybės vaisiai ir maišytos kokybės vaisiai. Duomenų rinkinys turi 19500 nuotraukų, kurių dydis 256x256 pikselių.

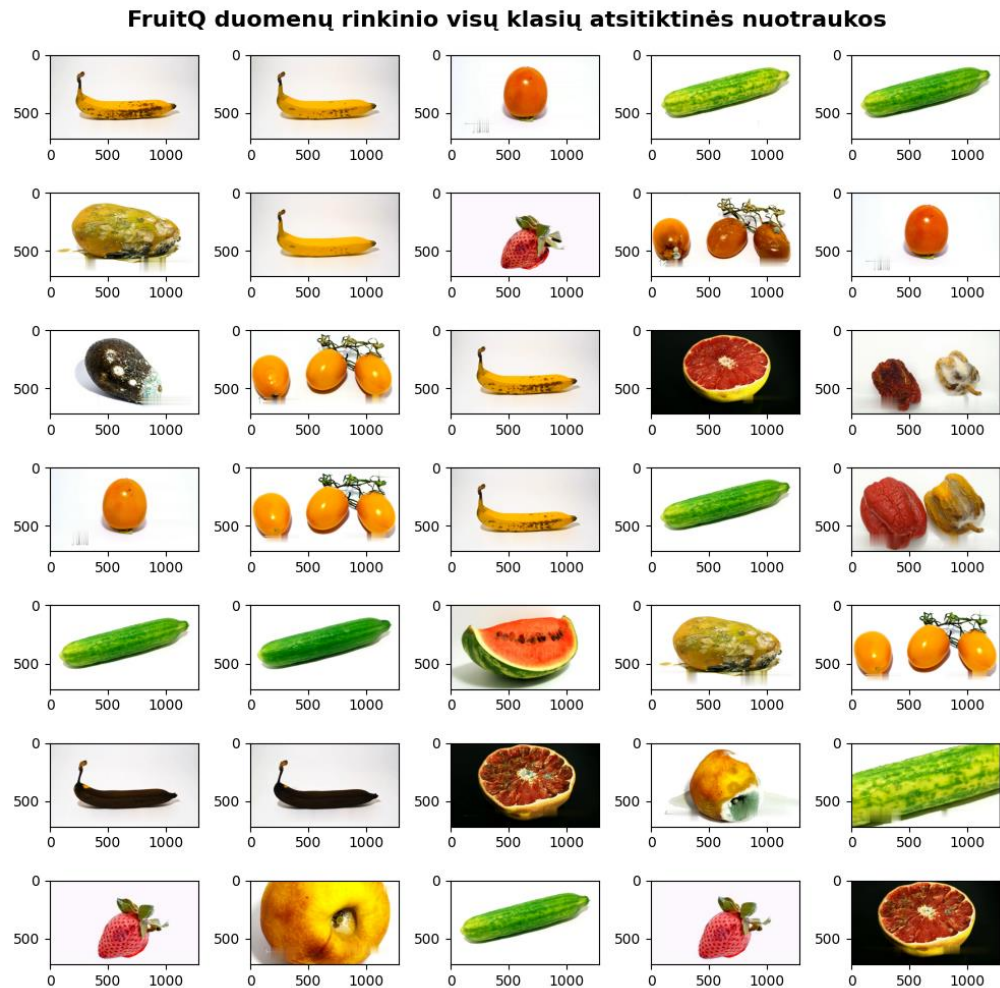
*Nusrat S. et al.* [25] 2022 metais pristatė šviežių ir supuvusių vaisių duomenų rinkinį, kuris buvo sudarytas iš nuotraukų darytų vaisių parduotuvėse ir realiuose vaisių laukuose. Duomenų rinkinį sudaro aštuonių rūšių vaisiai: vynuogės, gvajava, zizifas, granatas, obuolys, bananas, apelsinas, braškė. Rinkinį sudaro 12335 nuotraukos, kurių dydis 512x512 pikselių.

#### 2 lentelė. Vaisių duomenų rinkinių palyginimas

Duomenų rinkinys	Vaisių tipai	Nuotraukų kiekis	Klasių skaičius	Pritaikymo sritis
Fruits360	131	90380	131	Vaisių klasifikavimas
FruitNet	6	19500	6	Vaisių klasifikavimas Vaisių kokybės vertinimas
<i>Nusrat S. et al.</i> duomenų rinkinys	8	12335	16	Vaisių klasifikavimas Vaisių kokybės vertinimas
FruitQ	11	9421	33	Vaisių kokybės vertinimas

Vaisių aptikimui ir klasifikavimui buvo pasirinktas *Olusola O. Abayomi-Alli et al.* [26] pristatytas, atvirai prieinamas „FruitQ“ duomenų rinkinys. Duomenų rinkinys buvo sukurtas siekiant patobulinti išsamų vaisių atpažinimą ir jų kokybės nustatymą realiuoju laiku pagal konkrečius kriterijus, pavyzdžiui, spalvą, formą ir kt. „FruitQ“ duomenų rinkinys susideda iš nuotraukų sudarytų iš „YouTube“ vaizdo įrašų, kuriuose filmuojamas vaisių senėjimas. Šį duomenų rinkinį sudaro 11 vaisių

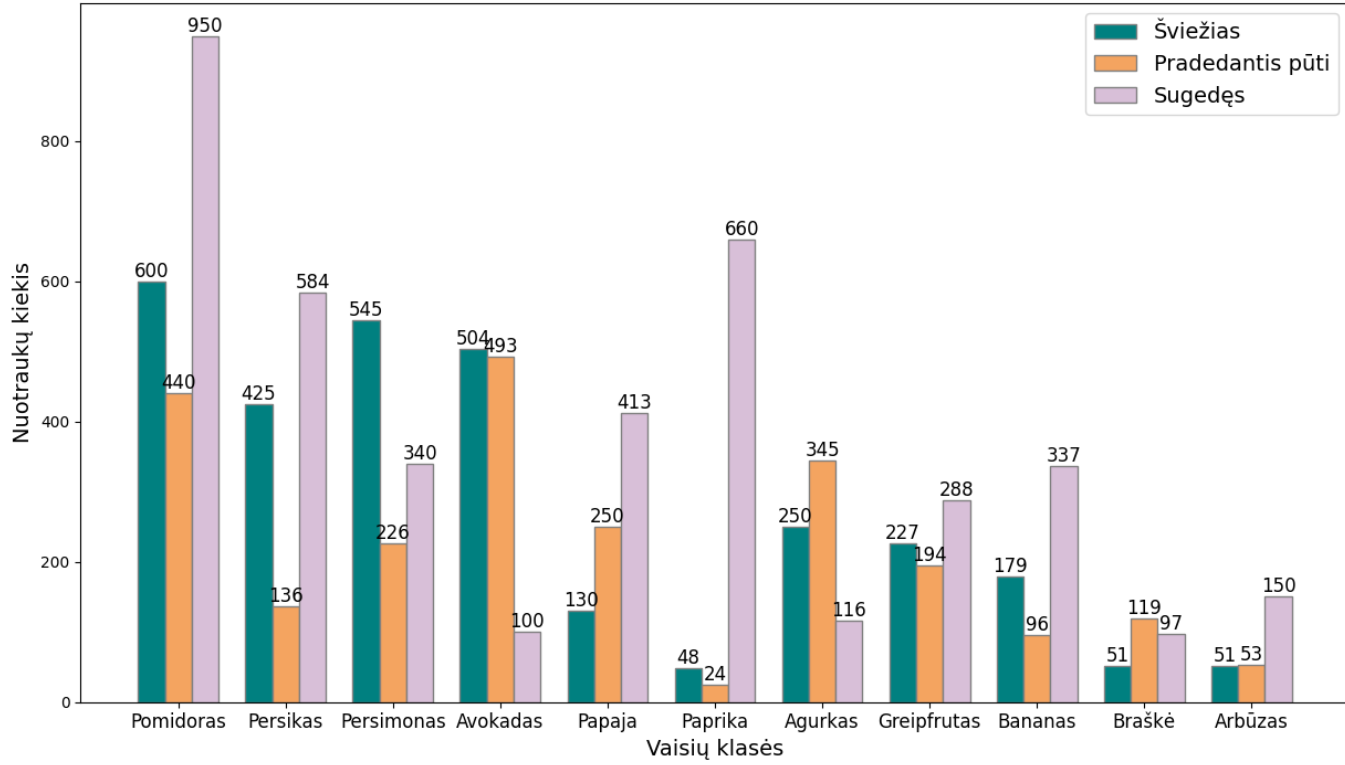
rūšių: bananas, agurkas, greipfrutas, persimonas, papaja, persikas, avokadas, paprika, braškė, pomidoras, arbūzas. Kiekvienas vaisius turi tris klases: šviežias, pradedantis pūti, supuvęs. Žemiau esančiame paveiksle atvaizduotos, atsitiktinai parinktos visų klasių „FruitQ“ duomenų rinkinio, nuotraukos.



**13 pav.** „FruitQ“ duomenų rinkinio atsitiktinai parinktos nuotraukos

Žemiau esančiame grafike atvaizduotas „FruitQ“ duomenų rinkinio pasiskirstymas, kuriame matomas kiekvieno vaisiaus nuotraukų kiekis pagal šviežumo klases.

**FruitQ duomenų rinkinio pasiskirstymas**

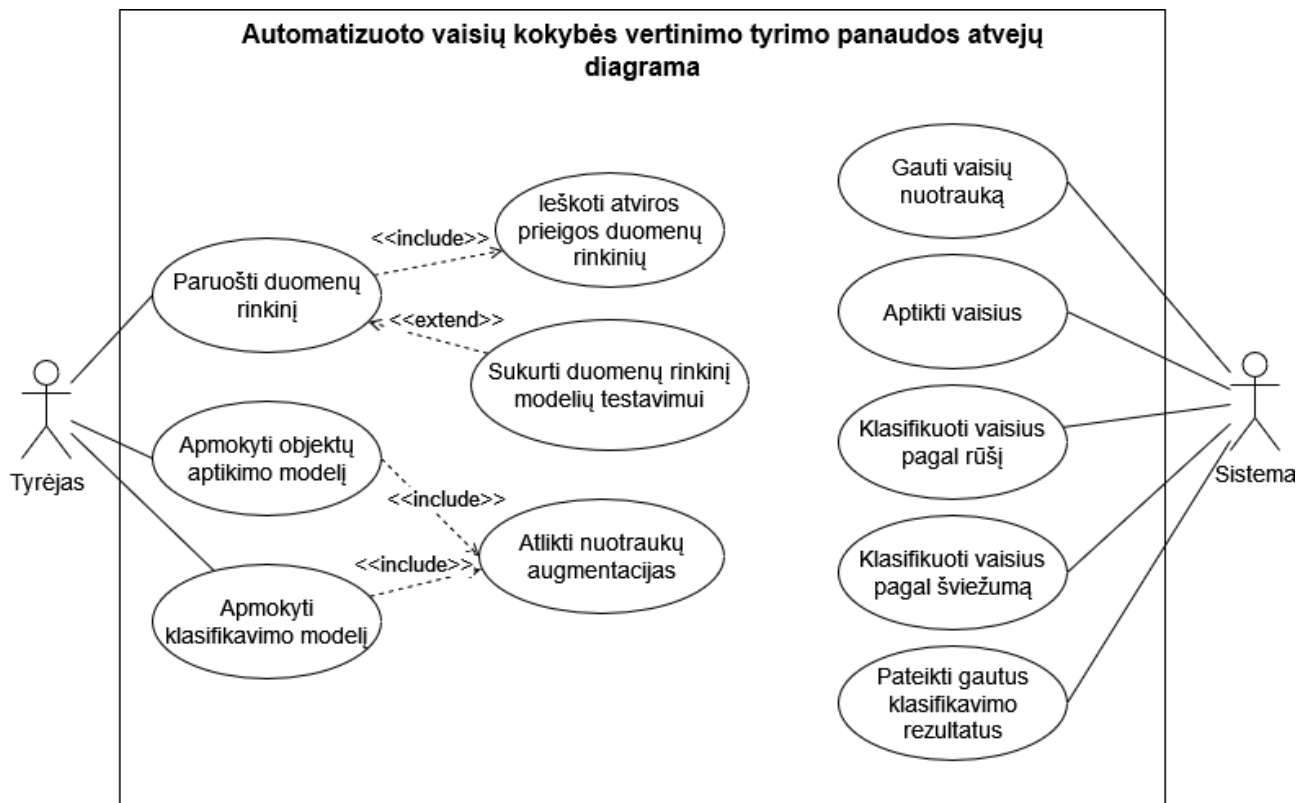


**14 pav.** „FruitQ“ duomenų rinkinio klasių pasiskirstymas

## 2. Automatizuoto vaisių kokybės tyrimo projektavimas

### 2.1. Sistemos koncepcija

Automatizuoto vaisių kokybės vertinimo tyrimas susideda iš tyrėjo, kuris atsakingas už duomenų paruošimą ir modelių apmokymą, ir sistemos, kuri veikia „Python“ kodo pavidalu.

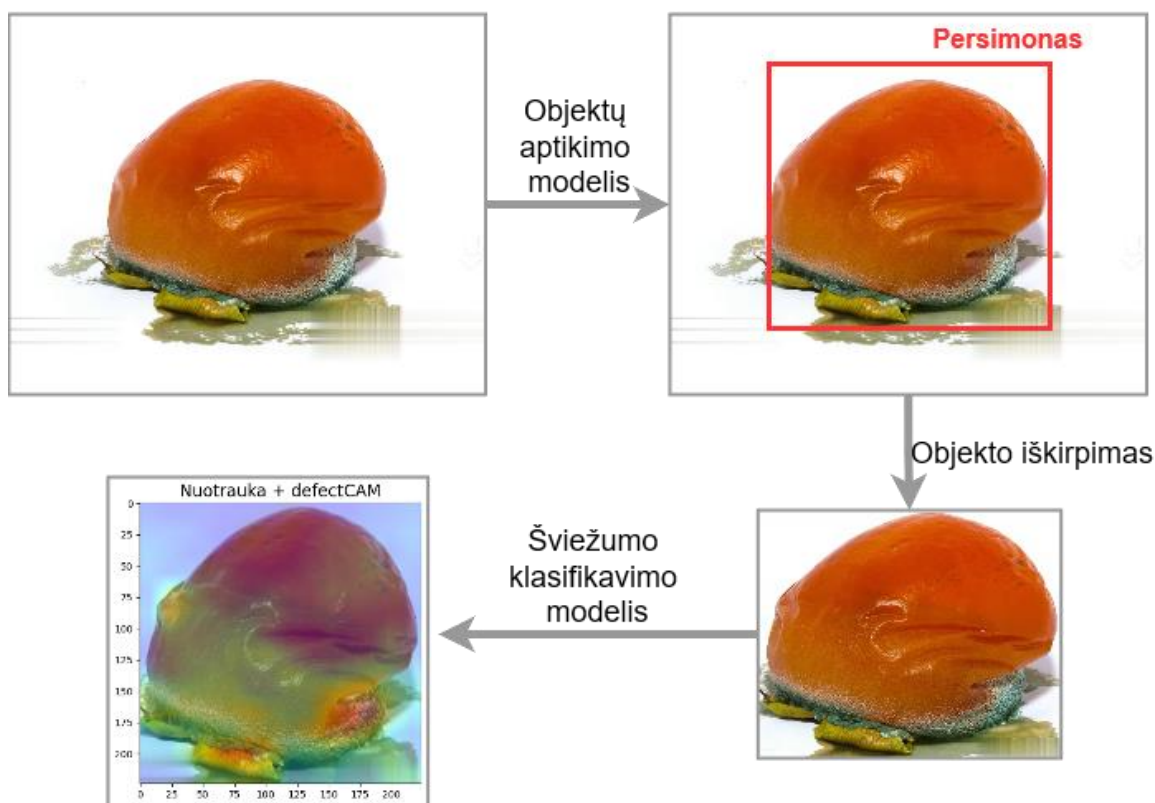


15 pav. Automatizuoto vaisių kokybės vertinimo tyrimo panaudos atvejų diagrama

Pirmu žingsniu tyrėjas atlieka literatūros analizę ir suranda atviros prieigos duomenų, skirtų vaisių kokybės vertinimui ir vaisių aptikimui bei klasifikavimui. Papildomai tyrėjas susikuria savo duomenų rinkinį, kuris bus panaudotas modelių rezultatų validumui testuoti. Tuomet duomenys sutvarkomi, atliekamos augmentacijos duomenų įvairovei bei kiekiui papildyti. Duomenys manualiai sužymimi objektų aptikimui.

Atlikus metodų ir architektūrų analizę išsirenkamos kelios modelių architektūros ir jos apmokomos su paruoštais duomenimis.

Sukurta sistema priima pavienes arba kelias nuotraukas, jos praleidžiamos pro apmokytą objektų aptikimo modelį. Aptiktas vaisius ar vaisiai yra iškerpami ir toliau iškarpa su prognozuota vaisiaus rūšimi siunčiama į atitinkamą kokybės klasifikavimo modelį. Galiausiai modelis grąžina rezultatą.



16 pav. Detalizuota sistemos koncepcija

## 2.2. Funciniai reikalavimai

Šiame poskyryje aprašomi funkciniai reikalavimai automatizuotai vaisių sistemai.

- Vartotojas gali įkelti pasirinktas nuotraukas.
- Sistema aptinka ir lokalizuoja vaisių esantį nuotraukoje, iškerpamas kadras ir grąžinama vaisiaus rūšis.
- Sistema atlieka vaisiaus kokybės klasifikavimą pagal gautą iškirptą kadra.
- Sistema grąžina vaisiaus kokybės rezultata.

## 2.3. Nefunkciniai reikalavimai

Šiame poskyryje aprašomi nefunkciniai reikalavimai automatizuotai vaisių sistemai.

- Duomenų rinkinyje kiekviena klasė turėtų turėti bent 300 nuotraukų.
- Naudojama „CVAT“ programa nuotraukų anotavimui.
- Kiekvieno vaisiaus rūšies suanotuota bent po 20 nuotraukų.
- „YOLOv8“ objektų aptikimo ir klasifikavimo tikslumas siekia bent 90%.
- Šviežumo klasifikavimo tikslumas siekia bent 90%.
- „Python>=3.10“ versija.
- „Tensorflow“ ir „Keras“ giliojo mokymosi modelių kūrimo karkasas.

## 2.4. Aparatūrinė įranga

Modelių mokymui naudojama „Google Colab“ aplinka, kuri nemokamai suteikia prieigą prie 16GB NVIDIA T4 vaizdo plokštės, 12GB atminties.

## 2.5. Sprendimo kūrimo priemonės

Sprendimui sukurti naudojama „Python“ atvirojo kodo programavimo kalba. Objektų aptikimo modelių mokymui naudojama „Ultralytics“ biblioteka, kuri leidžia išsikviesti „YOLOv8“ modelio architektūrą ir ją apmokyti. Kokybės klasifikavimo uždaviniams spręsti naudojama „Tensorflow“ [28] biblioteka, kuri leidžia importuoti „EfficientNet“, „MobileNet“ architektūras, bei suteikia įrankius realizuoti „Vision Transformer“ architektūrą. Nuotraukų augmentacijoms pritaikyti naudojama „Albumentations“ [27] biblioteka.

## 2.6. Vertinimo metrikos

Norint suprasti kaip gerai apmokytas modelis sugeba lokalizuoti vaisius bei juos klasifikuoti ir norint įvertinti modelio tikslumą bus apskaičiuojamos kelios dažniausiai naudojamos metrikos objektų aptikimo ir klasifikavimo uždaviniuose.

### 2.6.1. Tikslumas, preciziškumas, jautrumas

Norint įvertinti modelių tikslumą ir teisingumą, reikia sudaryti klaidų matricą (angl. *confusion matrix*). Sprendžiamas klasifikavimo uždavinys suskaidytas į dvi klases: 0 ir 1. Klasė „0“ žymi šviežius vaisius, o klasė „1“ – pradedančius pūti ar supuvusius vaisius. Klaidų matrica susideda iš keturių galimų kombinacijų:

- **Teisingi neigiami atvejai (TN):** Modelis teisingai prognozuoja klasę „0“ (faktinė reikšmė yra 0, prognozė yra 0).
- **Klaidingi teigiami atvejai (FP):** Modelis prognozuoja klasę „1“, nors teisinga klasė yra „0“ (faktinė reikšmė yra 0, prognozė yra 1).
- **Klaidingi neigiami atvejai (FN):** Modelis prognozuoja klasę „0“, nors teisinga klasė yra „1“ (faktinė reikšmė yra 1, prognozė yra 0).
- **Teisingi teigiami atvejai (TP):** Modelis teisingai prognozuoja klasę „1“ (faktinė reikšmė yra 1, prognozė yra 1).

Sudarius sumaišymo matricą, joje pateikiami šių kombinacijų atvejų skaičiai, lyginant tikrąsias faktines klases su modelio prognozėmis (17 pav.).

	Prognozuojama <b>0</b>	Prognozuojama <b>1</b>
Faktinė <b>0</b>	TN	FP
Faktinė <b>1</b>	FN	TP

17 pav. Sumaišymo matricos pavyzdys

Tikslumas (angl. *accuracy*) – teisingai suklasifikuotų klasių procentas.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Preciziškumas (angl. *precision*) – rodiklis, kuris nurodo, kiek iš visų klasės „1“ prognozių tikrai priklauso šiai klasei.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Jautrumas – (angl. *recall*) – rodiklis, kuris nurodo, kiek iš visų faktinių „1“ klasės reikšmių yra teisingai prognozuojama.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Įverčių metrikų reikšmių sritis yra režiuose [0; 1]. Kuo įvertis arčiau 1, tuo modelis tiksliau atlieka spėjimus.

### 2.6.2. mAP50 įvertis

Vertinant „YOLO“ modelio veikimo tikslumą dažniausiai matuojamos trys metrikos: vidutinio tikslumo vidurkis (angl. *Mean Average Precision* mAP), preciziškumas ir jautrumas. Įvertis mAP yra paskaičiuojamas sujungiant jautrumo ir preciziškumo gautas reikšmes.

Į šią metriką, taip pat, įtraukiamas ir aibių sankirtos bei sąjungos santykio (angl. *Intersection Over Union* IoU) vertė, kuri paskaičiuojama taip:

$$IoU = \frac{Sankirtos\ plotas}{Sajungos\ plotas} \quad (4)$$

Čia imamas faktinių objektą apibrėžiančių koordinatų plotas ir modelio prognozuojamų objektą apibrėžiančių koordinatų plotas.

Tuomet mAP50 įverčiui paskaičiuojamas vidutinio tikslumo vidurkis, kuomet IoU reikšmė pasiekia 0,5. Tai reiškia jog modelio prognozuojamas aptikimas skaitomas, kaip teisingas jei faktinių objektą apibrėžiančių koordinatų ir modelio prognozuojamų koordinatų IoU įvertis peržengia 0,5.

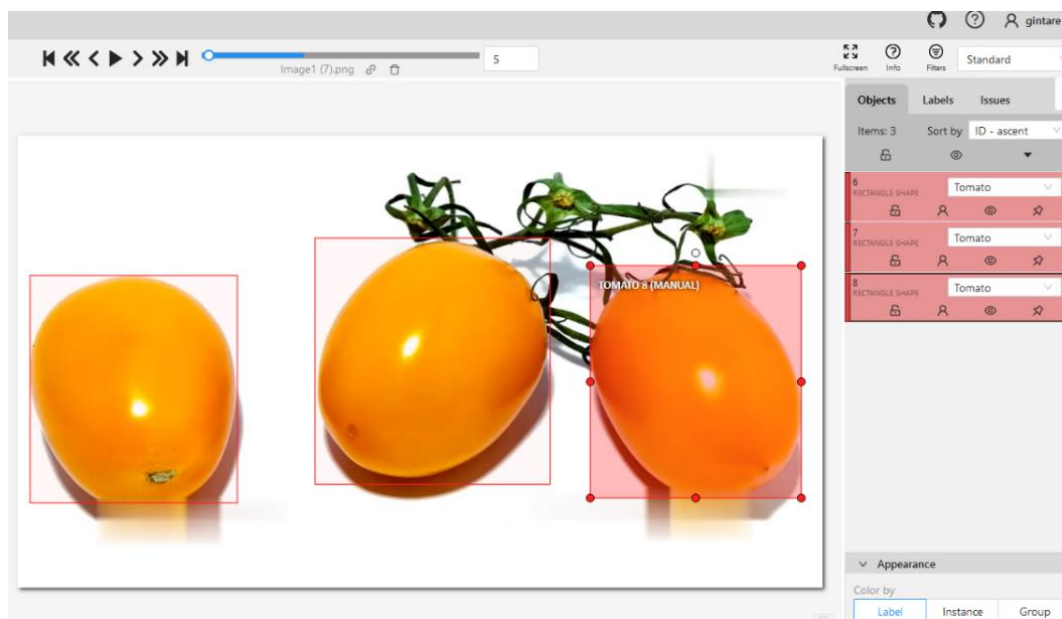
### 3. Eksperimentinė dalis

#### 3.1. Duomenų paruošimas

Iš *FruitQ* duomenų rinkinio buvo atmesta braškės klasė, kadangi parsisiųstame faile ši klasė turėjo tik šviežių nuotraukų katalogą. Taip pat, klasifikavimo modelių mokymui atsisakyta arbūzo klasė, kadangi net šviežios klasės kataloge, šis vaisius buvo pradėjęs pūti. Toliau objektų aptikimo ir klasifikavimo modeliui buvo naudojamos visos likusios klasės: agurkas, avokadas, bananas, greipfrutas, papaja, paprika, persikas, persimonas ir pomidoras. Visos vaisių klasės surūšiuotos į dvi klases: šviežias vaisius ir pradėjęs pūti vaisius. Kiekvienai klasei parinkta 80 nuotraukų, taip užtikrinant subalansuotą duomenų rinkinį. Kiekviena klasė suskaidyta į mokymo ir validavimo imtis kuomet 80% duomenų naudojame mokymui, o likusius 20% - testavimui. Atlikus šį skaidymą nuotraukų kiekis kiekvienai klasei pasiskirstė: 64 nuotraukų modelio mokymui, 16 nuotraukų validavimui.

##### 3.1.1. Duomenų paruošimas objektų aptikimo uždaviniui

Objektų atpažinimo užduočiai spręsti duomenys buvo pasiruošti naudojant nemokama „CVAT“ [29] duomenų žymėjimo programa. Iš kiekvienos klasės pasirinkta 20 vaisių nuotraukų įtraukiant šviežius, pradedančius pūti ir supuvusius vaisius. Šios nuotraukos rankiniu būdu sužymėtos kvadratėliu apvedant vaisių ir parenkant atitinkamą klasę. Žymėjimai išeksportuoti YOLO formatu.



18 pav. „CVAT“ programos duomenų anotavimo langas

##### 3.1.2. Duomenų paruošimas kokybės klasifikavimo uždaviniui

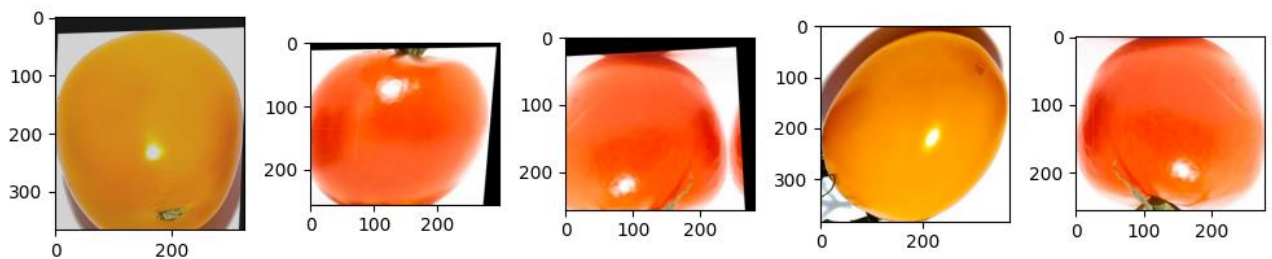
Vaisių kokybės klasifikavimo uždaviniui iš anksto pasiruoštos nuotraukos su pritaikytomis augmentacijomis. Augmentacijų pritaikymui naudojama „Albumentations“ biblioteka, 3 lentelėje aprašytos naudotos augmentacijos bei kiekvienos iš jų parametrai.

3 lentelė. Naudotų augmentacijų parametrai

<i>Albumentations</i> augmentacija	Augmentacijos parametrai
VerticalFlip – vertikalus apvertimas	p=0.5
HorizontalFlip – horizontalus apvertimas	p=0.5
Sharpen – paryškimas	p=0.5
ShiftScaleRotate – vaizdo pritraukimas/atitolinimas, vaizdo pasukimas	p=0.3, shift_limit=[-0.1, 0.1], scale_limit=[-0.05, 0.05], rotate_limit=[-5, 5]
Affine – vaizdo pastūmimas x/y ašimis	shear=(-5, 5), translate_percent={"x": (-0.1, 0.1), "y": (-0.1, 0.1)}

Kiekvienai nuotraukai, esančiai mokymo imtyje, pritaikytos penkios augmentacijos iš aukščiau paminėto sąrašo. Atlikus duomenų augmentaciją kiekviena klasė turėjo po 384 nuotraukas. Žemiau esančiame 19 paveiksle atvaizduotos atsitiktinai parinktos augmenteduotos nuotraukos.

**Atliktų augmentacijų pavyzdinės nuotraukos**

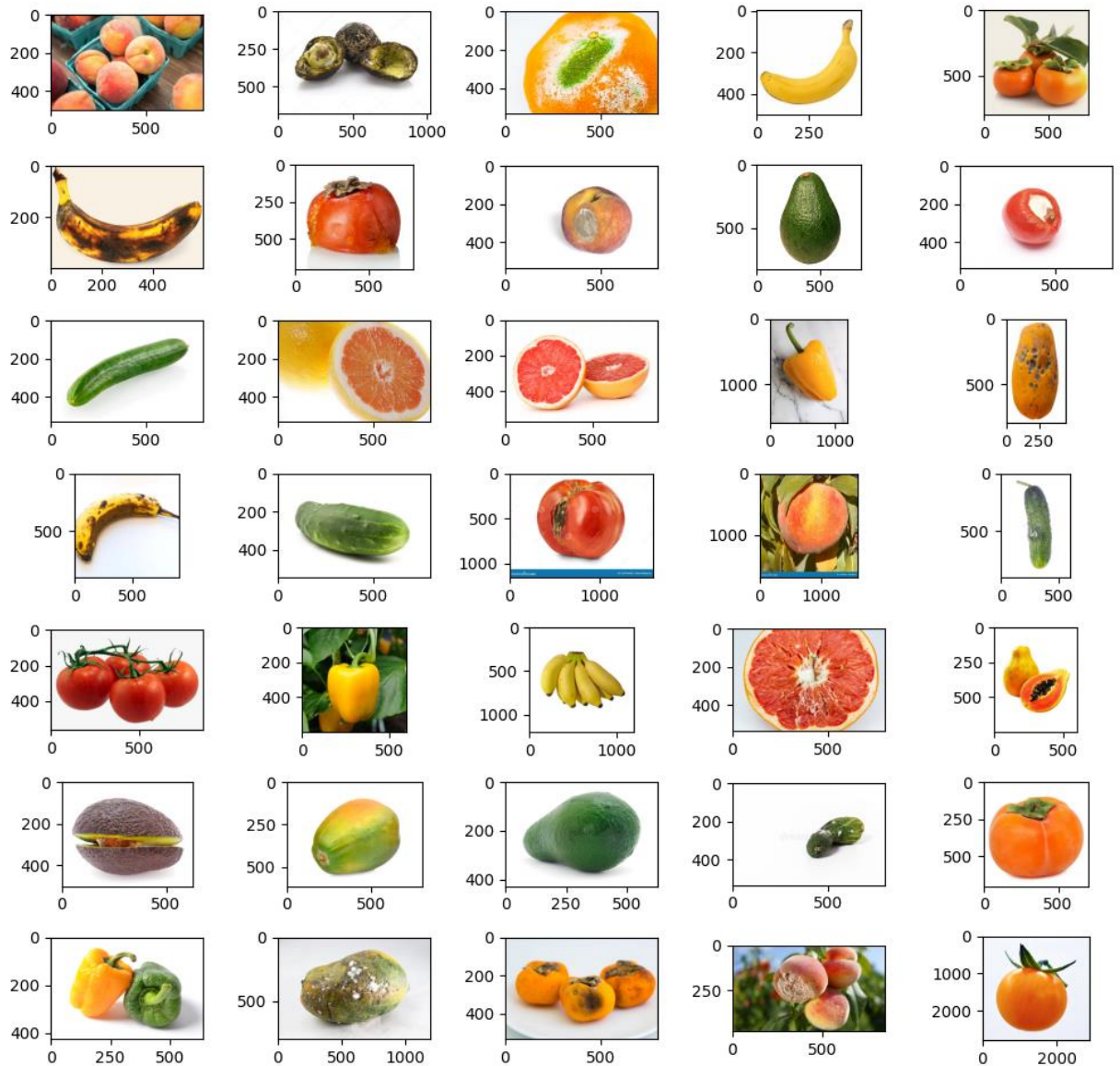


**19 pav.** Pritaikytų augmentacijų pavyzdžiai

### 3.1.3. Duomenų paruošimas modelių rezultatų testavimui su nematytais duomenimis

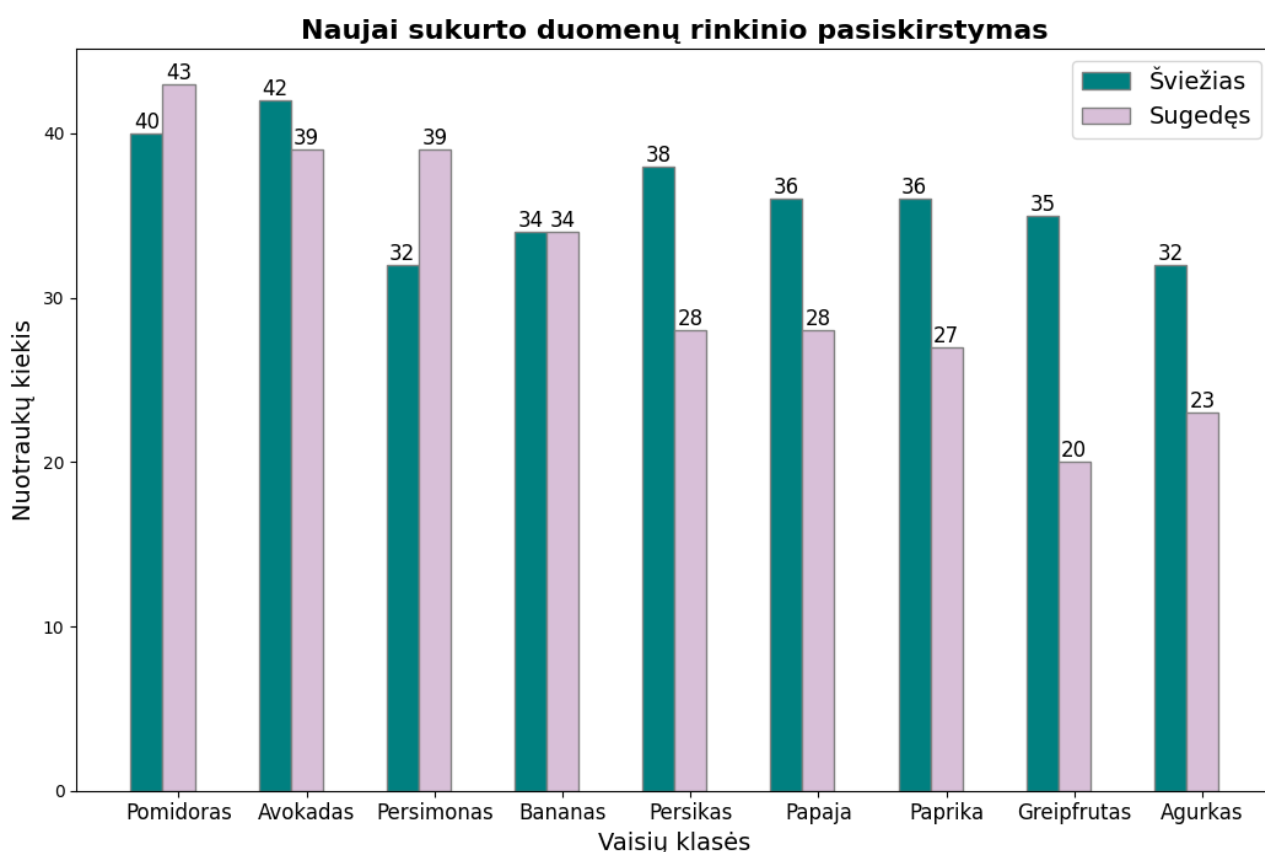
Modelių validumui įrodyti, bei įvertinti, „FruitQ“ atvirojo duomenų rinkinio, pritaikomumą vaisių kokybės vertinimui, sukurtas nuosavas duomenų rinkinys. Šis duomenų rinkinys susideda iš 9 vaisių klasių, kurios suskirstytos į „šviežius“ ir „sugedusius“ vaisius. Vaisių rūšys atitinka „FruitQ“ vaisių rūšis. Duomenų rinkinys sudarytas iš „Bing“ naršyklės nuotraukų renkant jas automatinio būdu. Surinktos nuotraukos vėliau manualiniu būdu peržiūrėtos, ir blogos kokybės ar netinkamos nuotraukos atmestos. Nuotraukų dydžiai varijuojantys. Šis duomenų rinkinys gali būti pritaikomas vaisių rūšių klasifikavimui bei vaisių kokybės nustatymui. Žemiau esančiame paveiksle (20 pav.) atsitiktiniu būdu parinktos iš atvaizduotos šio duomenų rinkinio nuotraukos.

### Naujai sukurto duomenų rinkinio visų klasių atsitiktinės nuotraukos



**20 pav.** Naujai sukurto duomenų rinkinio visų klasių atsitiktinai parinktos nuotraukos

Šį duomenų rinkinį sudaro 606 originalios nuotraukos be augmentacijų. Žemiau esančiame paveiksle (21 pav.) pateikiamas kiekvienos vaisių klasės nuotraukų kiekis.



**21 pav.** Naujai sukurto duomenų rinkinio pasiskirstymas

### 3.2. Atrinktų modelių mokymas

Vaisių aptikimui ir rūšies nustatymui buvo apmokytos 4 modelio versijos, pasinaudojant anksčiau sužymėtais duomenimis. Kiekvienai modelio versijai naudota ta pati duomenų imtis. Naudotos numatytosios „YOLOv8“ modelio augmentacijos. Versijos skiriasi tuom, jog naudotos skirtingos *dropout* reikšmės, jas didinant. Taip pat, paskutinio bandymo epochų skaičius padidintas, prailginant mokymą dar papildomų 20 epochų. 4 lentelėje aprašyti bandymuose keisti parametrai.

**4 lentelė.** „YOLOv8“ skirtingų modelio versijų hiperparametrai

Modelio versija	Epochų skaičius	<i>Dropout</i> reikšmė	Vaizdo dydis, px	Pradinė mokymosi reikšmė	Optimizavimo funkcija
v1.0	30	0,0	640x640	0,01	Adam
v2.0	30	0,1	640x640	0,01	Adam
v3.0	30	0,2	640x640	0,01	Adam
v4.0	50	0,2	640x640	0,01	Adam

Vaisių kokybės vertinimui apmokytos 4 skirtingos modelių architektūros: „EfficientNetV2B0“, „EfficientNetV2B3“, „MobileNetV3“, „Vision Transformer“. Modelių mokymui naudotas paruoštas „FruitQ“ duomenų rinkinys. Pirmieji du bandymai atlikti apmokant duomenis, kuriems nebuvo pritaikyta jokia augmentacija. Sekantiems bandymams pritaikyti anksčiau augmentuoti duomenys.

Įvesties vaizdo dydis parinktas kiekvienai architektūrai individualiai, pagal „Tensorflow“ dokumentacijos rekomendacijas. Epochų skaičius parinktas pagal atliktus bandymus, kadangi pastebėta, jog po 20 epochų modelių rezultatai nusistovi ir nekinta. Dėl architektūrinių skirtumų „Vision Transformer“ modeliui epochų skaičius padvigubintas, kadangi ties 20 epochų rezultatų nusistovėjimas nepastebėtas. Papildomi bandymai atlikti su „EfficientNetV2B0“ architektūra įtraukiant papildomą *dropout* sluoksnį su 3 skirtingomis tikimybinėmis reikšmėmis: 0,2, 0,3, 0,4. Šis sluoksnis įtrauktas su tikslu mažinti modelio persimokymą. Šiam papildomam bandymui pasirinkta „EfficientNetV2B0“ architektūra dėl aukštų metriku, kurios aptariamose 3.3. modelių rezultatų analizės skyrelyje. Visi keisti hiperparametrai bandymų metu pateikti 5 lentelėje.

**5 lentelė.** Kokybės klasifikavimo modelių hiperparametrai

<b>Modelis</b>	<b>Epochų skaičius</b>	<b>Dropout reikšmė</b>	<b>Vaizdo dydis, px</b>	<b>Pradinė mokymosi reikšmė</b>	<b>Optimizavimo funkcija</b>
EfficientNetV2B0	20	0,0	224x224	0,001	Adam
MobileNetV3	20	0,0	224x224	0,001	Adam
EfficientNetV2B0 + augmentacijos	20	0,0	224x224	0,001	Adam
EfficientNetV2B3 + augmentacijos	20	0,0	300x300	0,001	Adam
MobileNetV3 + augmentacijos	20	0,0	224x224	0,001	Adam
EfficientNetV2B0 + aug. + 0,2 <i>dropout</i>	20	0,2	224x224	0,001	Adam
EfficientNetV2B0 + aug. + 0,3 <i>dropout</i>	20	0,3	224x224	0,001	Adam
EfficientNetV2B0 + aug. + 0,4 <i>dropout</i>	20	0,4	224x224	0,001	Adam
ViT + augmentacijos	40	0,0	224x224	0,001	AdamW

Transformerių modeliui apmokyti buvo naudojami šie hiperparametrai, aprašyti 6 lentelėje.

**6 lentelė.** Transformerių modelio hiperparametrai

<b>Parametras</b>	<b>Reikšmė</b>
Karpomų dalių dydis	16x16 pikselių
Karpomų dalių projekcijos dydis	64x64 pikselių
Transformerių sluoksnių kiekis	4 sluoksniai
Dėmesio galvų kiekis	4 galvos
Transformerio sujungtų sluoksnių kiekis	2 sluoksniai
Transformerio sujungtų sluoksnių dydis	128 neuronai
Visiškai sujungtų sluoksnių kiekis	2 sluoksniai
Visiškai sujungtų sluoksnių dydis	1024 neuronų

### **3.3. Modelių rezultatų analizė**

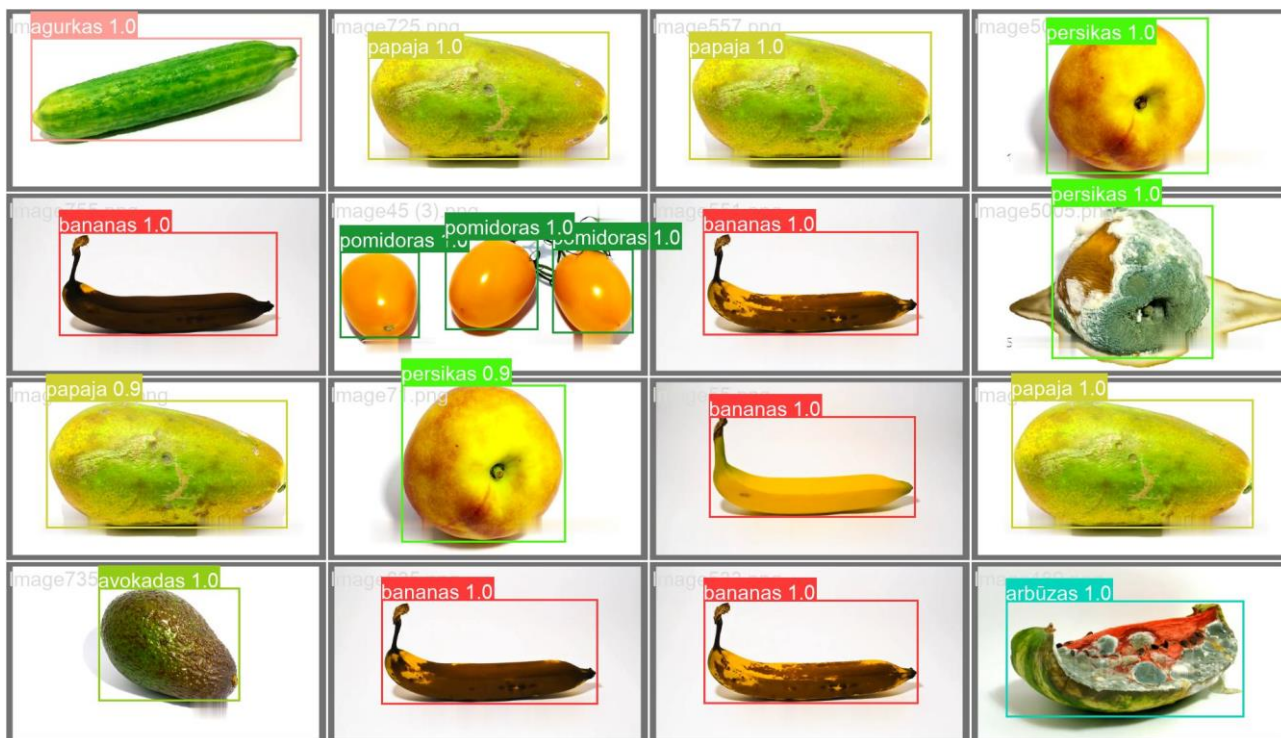
Objektų aptikimo ir klasifikavimo modelio mokymo metu gauti tikslumo įverčiai pateikti 7 lentelėje. Aukščiausią įvertį pasiekė v4.0 „YOLOv8“ modelio versija, pasiekusi 99,3% mAP50 metrikos tikslumo.

*Precision* įvertis, taip pat, rodo jog modelis 98,12% tikslumu klasifikuoja vaisius pagal jų rūšį. v4.0 modelis pasiekė 0,36% geresnius rezultatus nei v3.0 modelio versija, kuomet epochų skaičius padidintas papildomų 20 epochų.

**7 lentelė.** „YOLOv8“ skirtingų modelių mokymo rezultatai

<b>Modelio versija</b>	<b>mAP50 metrikos reikšmė</b>	<b><i>Precision</i> reikšmė</b>
v1.0	0,9856	0,9569
v2.0	0,9862	0,9724
v3.0	0,9884	0,9744
v4.0	0,9929	0,9812

Geriausio modelio spėjimai ant validavimo duomenų atvaizduoti 22 pav., iš pateiktų rezultatų matome, jog modelis teisingai aptinka visus objektus matomus nuotraukoje ir klasifikuoja juos 90-100% rėžių tikslumu.



22 pav. Geriausio „YOLOv8“ modelio rezultatai

8 lentelėje pateikti 3.2 skyrelyje aprašytų modelių bandymų rezultatai. Prasčiausiai pasirodė modeliai, kurių mokymui buvo naudojami duomenys be papildomų augmentacijų. Tokį rezultatą galėjo lemti maža įvairovė skirtumų tarp vaisių esančių toje pačioje klasėje.

Geresni rezultatai pastebimi bandymų, kuriuose buvo naudotos augmentuotos nuotraukos. Tačiau, net su augmentacijomis „MobileNetV3“ architektūra nepranoko „EfficientNetV2B0“ gautų rezultatų. Taip pat, „EfficientNetV2“ architektūros „B3“ versija nesugebėjo gauti geresnių rezultatų nei „B0“ versija. Tokį rezultatą galėjo lemti naudotas vienodas epochų skaičius, kadangi „B3“ versijos architektūra turi daugiau parametrų, todėl gali prireikti didesnio kiekio iteracijų išmokti skirtumus tarp skirtingų klasių.

8 lentelė. Kokybės klasifikavimo modelių rezultatai

Modelis	Accuracy reikšmė	Recall reikšmė	Precision reikšmė
EfficientNetV2B0	0,9346	0,9145	0,9394
MobileNetV3	0,9184	0,8934	0,9023
EfficientNetV2B0 + augmentacijos	0,9931	0,9997	0,9994
EfficientNetV2B3 + augmentacijos	0,9651	0,9786	0,9679
MobileNetV3 + augmentacijos	0,9321	0,9444	0,9605
<b>EfficientNetV2B0 aug. + 0,2 dropout</b>	<b>0,9943</b>	<b>1,000</b>	<b>0,9979</b>
EfficientNetV2B0 + aug. + 0,3 dropout	0,9927	0,9927	0,9927

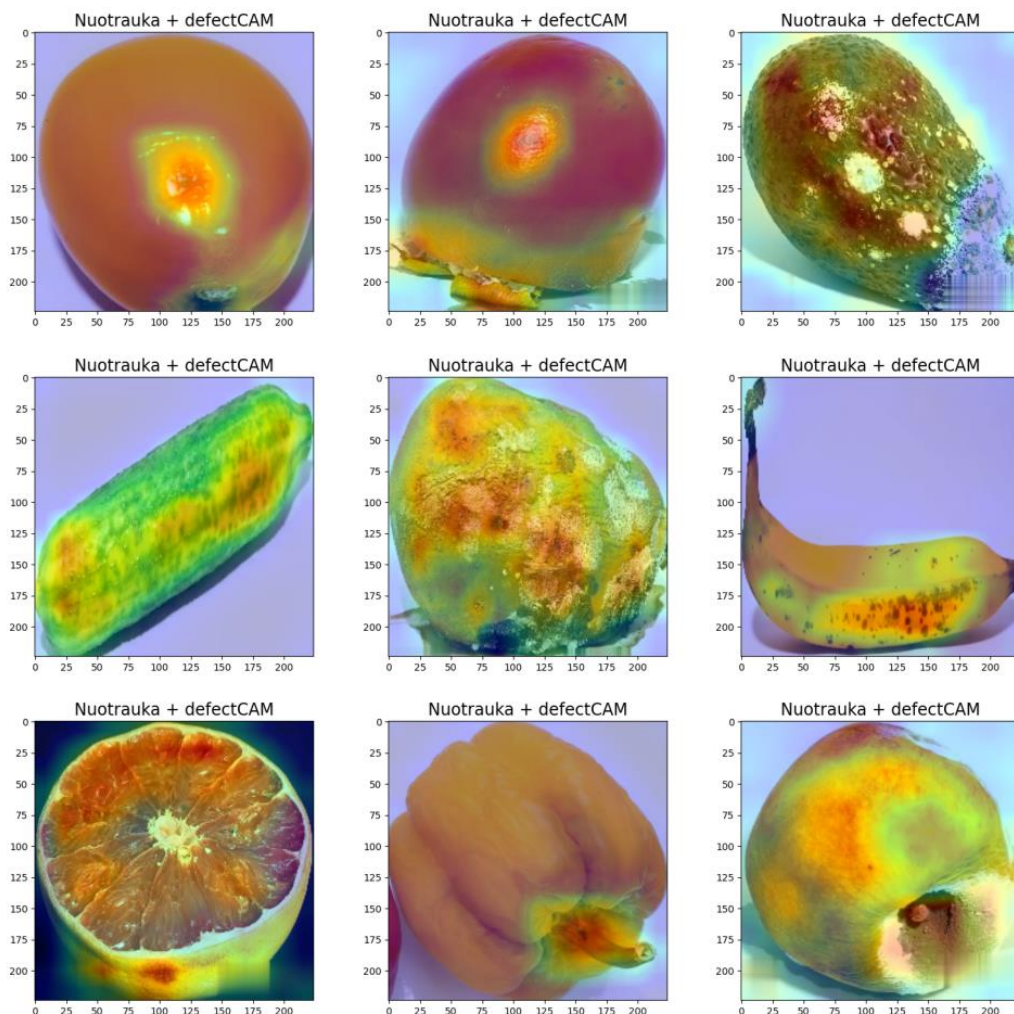
EfficientNetV2B0 + aug. + 0,4 dropout	0,9921	0,9918	0,9975
ViT + augmentacijos	0,9359	0,9549	0,9452

Žemiau esančiame 23 pav. atvaizduotos, geriausias rezultatus pasiekusio modelio, atsitiktinės nuotraukos, kuriose pritaikytas Grad-CAM. Šis atvaizdavimo metodas leidžia geriau suprasti modelio daromus sprendimus, kadangi išryškina vietas į kurias reaguoja modelis priimdamas sprendimą. Šiuo atveju matome, jog modelis gan teisingai supranta, kur vyksta vaisiaus pūvimas ir reaguoja į pradėjusias pūti vietas.

Pažvelgus į avokado ir paprikos klasės nuotraukas, galime pastebėti, jog modelis ne visai teisingai žymi pradėjusias gesti zonas. Didžioji dalis paprikos yra pradėjusi pūti, tačiau modelis reaguoja tik į kotelį, kuris neturi tiek daug reikšmingos informacijos, kiek pasikeitusi paprikos spalva kitose zonose.

Taip pat, avokado nuotraukoje modelis neišryškina būtent pūvimo vietų (baltų pelyjančių taškelių), o reaguoja į spalvos pokytį, kurį įgauna avokadas labiau sunokęs.

**Geriausio modelio reagavimo zonos kiekvienai klasei**



**23 pav. „EfficientNetV2B0“ modelio rezultatų reagavimo zonos**

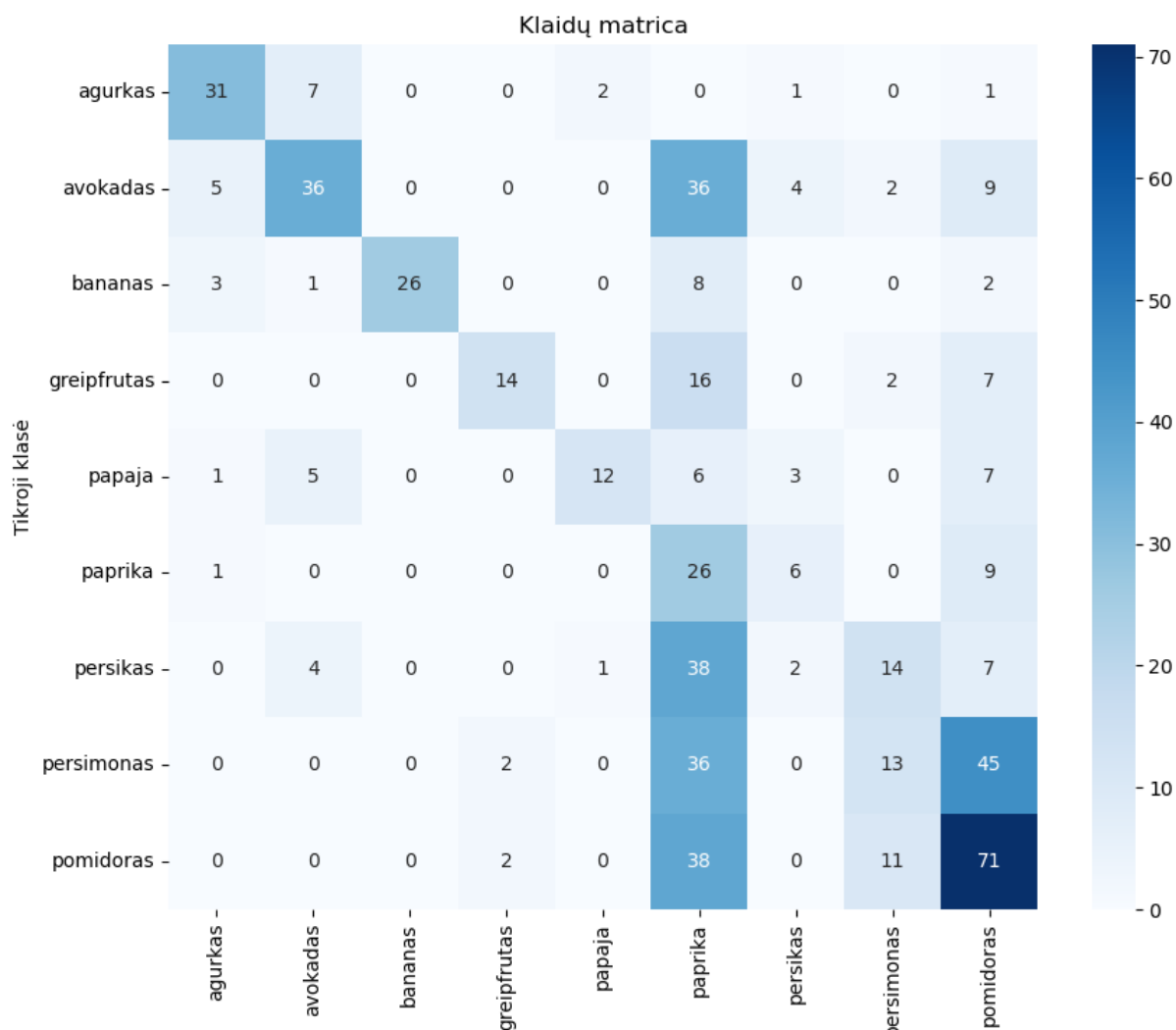
### 3.4. Modelių pritaikomumo naujiems duomenų rinkiniams įvertinimas

Jog įvertinti sistemos patikimumą ir gebėjimą veikti įvairesnėmis sąlygomis atliktas bandymas ant modeliui dar nematytų duomenų. Tam tikslui naudojamas 3.1.3 skyrelyje aprašytas nuosavas duomenų rinkinys.

#### 3.4.1. Vaisių aptikimo ir rūšies klasifikavimo tikslumas

Testavimui naudojama apmokyta „YOLOv8“ architektūros v4.0 modelio versija, kuri „FruitQ“ duomenų rinkiniui pasiekė 99,3% objektų aptikimo ir klasifikavimo tikslumą.

Tačiau su naujuoju duomenų rinkiniu pasiektas tikslumas gan žemas – 40%. Tokį mažą tikslumą galėjo lemti maža vaisių bei fono įvairovė „FruitQ“ duomenų rinkinyje. Nuosavame duomenų rinkinyje ne visi vaisiai buvo tik baltame fone, taip pat, dažnai vaisiai nuotraukose būna po kelis, o ne po vieną.



24 pav. v4.0 versijos objektų aptikimo modelio klaidų matrica su nematytais duomenimis

Paanalizavus klaidų matricą galima pastebėti, kur modelis klydo labiausiai. Eilutės simbolizuoja faktines klases, o stulpeliai – modelio spėjimus. Iš klaidų matricos matoma, jog gan dažnai klases: persikas, persimonas, pomidoras bei paprika maišomos tarpusavyje.

Pažvelgus į anksčiau aprašytus duomenų rinkinius ir jų nuotraukas, galime susidaryti spėjimą, jog šių klasių maišymą lemia šių vaisių dominuojanti spalva bei forma. Visi šie vaisiai turi dominuojančią spalvą raudoną/oranžinę, taip pat, yra apvalios formos. Taikant įvairias pasukimo, pavertimo augmentacijas, dėl apvalios formos nelabai matomas skirtumas, nes forma išieka visais atvejais vienoda.

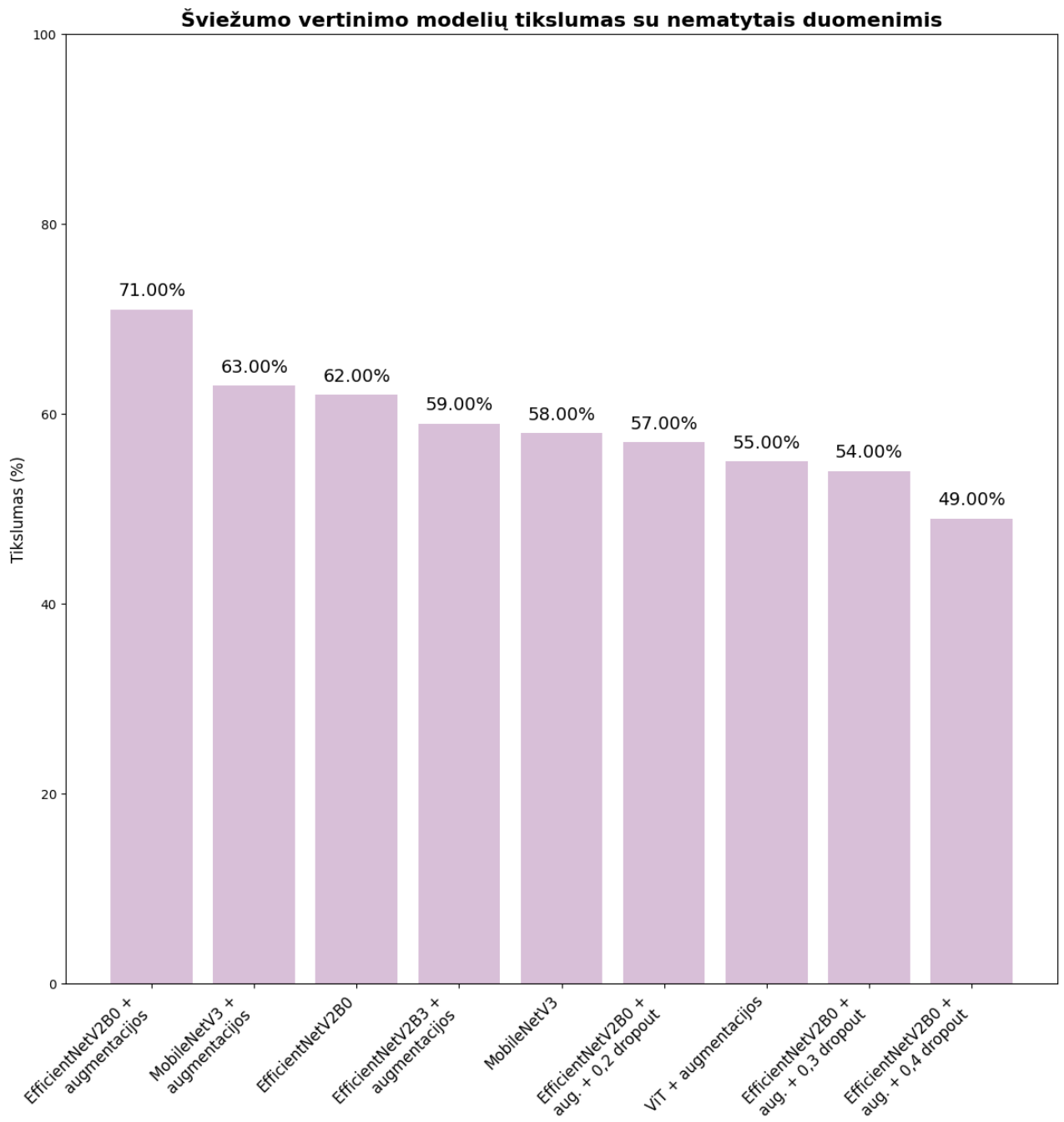
### 3.4.2. Vaisių kokybės vertinimo tikslumas

Vaisiai, kurie praeitame žingsnyje buvo aptikti, iškirpti ir suklasifikuoti pagal rūšį toliau perduoti atitinkamam šviežumo klasifikavimo modeliui.

9 lentelėje pateikti visų anksčiau apmokytų modelių tikslumo rezultatai naujam duomenų rinkiniui. Gauti rezultatai, taip pat, nėra labai aukšti. Geriausiai pasirodęs modelis buvo „EfficientNetV2B0“ su augmentacijomis, kuris pasiekė 71% klasifikavimo tikslumą. „FruitQ“ duomenų rinkinio klasifikavime geriausiai pasirodęs modelis „EfficientNetV2B0“ su augmentacijomis ir 0,2 *dropout* reikšme, naujame duomenų rinkinyje pasirodė su vienu iš žemiausių tikslumu – 54%.

**9 lentelė.** Vaisių šviežumo vertinimo sistemos tikslumas su nematytais duomenimis

Modelis	<i>Accuracy</i> reikšmė	<i>Recall</i> reikšmė	<i>Precision</i> reikšmė
EfficientNetV2B0	0,62	0,56	0,63
MobileNetV3	0,58	0,57	0,34
<b>EfficientNetV2B0 + augmentacijos</b>	<b>0,71</b>	<b>0,71</b>	<b>0,74</b>
EfficientNetV2B3 + augmentacijos	0,59	0,53	0,56
MobileNetV3 + augmentacijos	0,63	0,56	0,81
EfficientNetV2B0 + aug. + 0,2 <i>dropout</i>	0,57	0,57	0,54
EfficientNetV2B0 + aug. + 0,3 <i>dropout</i>	0,54	0,54	0,50
EfficientNetV2B0 + aug. + 0,4 <i>dropout</i>	0,49	0,45	0,44
ViT + augmentacijos	0,55	0,55	0,53



**25 pav.** Vaisių šviežumo vertinimo sistemos tikslumas su nematytais duomenimis

## Išvados

1. Literatūros analizės metu buvo apžvelgtos skirtingos modelių architektūros, naudojamos vaisių rūšies klasifikavimo bei vaisių kokybės vertinimo uždaviniams spręsti. Remiantis analizės rezultatais, eksperimentai buvo atlikti apmokant 4 skirtingų architektūrų modelius vaisių kokybės klasifikavimui ir 1 architektūros modelis vaisių aptikimui ir rūšies klasifikavimui.
2. Pasirinktos modelių architektūros buvo apmokytos naudojantis vieštai prieinamu „FruitQ“ duomenų rinkiniu, pritaikytu vaisių kokybės vertinimo uždaviniams. Atlikus vaisių aptikimo ir rūšies klasifikavimo modelio mokymus, geriausius rezultatus pasiekė v4.0 „YOLOv8“ modelio bandymas, pasiekęs 99,3% mAP50 įvertį. Atlikus vaisių kokybės vertinimą buvo nustatyta, jog geriausius rezultatus pasiekė „EfficientNetV2B0“ architektūros modelis, kurio tikslumas siekia 99,3%, o prasčiausius rezultatus pasiekė „MobileNetV3“ architektūros modelis, kurio tikslumas siekia 91,8%.
3. Vaisių šviežumo klasifikavimo modeliams buvo pritaikytos, savo nuožiūra parinktos, augmentacijos, kurios pagerino modelio „EfficientNetV2B0“ tikslumą. Be augmentacijų modelio tikslumas siekia tik 93,5%. Taip pat, pridėtas papildomas duomenų išmetimo sluoksnis, kuris atsitiktinai išmeta 20% duomenų, su tikslu mažinti persimokymą, tai leido modeliui pasiekti 99,4% tikslumą.
4. Sukurtas nuosavas duomenų rinkinys, kuris skirtas vaisių aptikimo ir klasifikavimo uždaviniams, bei vaisių šviežumo vertinimo uždaviniams. Pritaikius anksčiau mokytus modelius šiem duomenims gauti tokie rezultatai: 40% vaisių aptikimo ir rūšies klasifikavimo tikslumas, ir 71% vaisių šviežumo klasifikavimo tikslumas.

## Literatūros sąrašas

1. Nazrul I., Owais A. Malik. *Real-time visual inspection system for grading fruits using computer vision and deep learning techniques*. Information processing in agriculture, 2022. Prieiga per internetą: <https://doi.org/10.1016/j.inpa.2021.01.005>
2. Rohini C, Geetha PS, Vijayalakshmi R, Mini ML and Pasupathi E. *Global effects of food waste*. J Pharmacogn Phytochem, 2020. Prieiga per internetą: <https://www.phytojournal.com/archives/2020.v9.i2.10936/global-effects-of-food-waste>
3. T. Bharath Kumar, Deepak Prashar, Gayatri Vaidya, Vipin Kumar, S. Deva Kumar, F. Sammy. *A Novel Model to Detect and Classify Fresh and Damaged Fruits to Reduce Food Waste Using a Deep Learning Technique*. 2022. Prieiga per internetą: <https://www.hindawi.com/journals/jfq/2022/4661108/>
4. Jayanth, J., Mahadevaswamy, M., Shivakumar, M. *Fruit Quality Identification and Classification by Convolutional Neural Network*. SN COMPUT. SCI., 2023. Prieiga per internetą: <https://doi.org/10.1007/s42979-022-01527-w>
5. Dhiman B., Kumar Y., Kumar M. *Fruit quality evaluation using machine learning techniques: review, motivation and future perspectives*. Multimedia Tools and Applications, 2022. Prieiga per internetą: 10.1007/s11042-022-12652-2
6. K. B. Mirra, R. Rajakumari. *Fruit quality detection and classification: A survey*. AIP Conf. Proc., 2022. Prieiga per internetą: <https://doi.org/10.1063/5.0078329>
7. *Convolutional Neural Networks, Explained*, 2020. Prieiga per internetą: <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>
8. *Activation Functions in Neural Networks*, 2017. Prieiga per internetą: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>
9. *Convolutional Neural Networks (CNNs) and Layer Types.*, 2021. Prieiga per internetą: <https://pyimagesearch.com/2021/05/14/convolutional-neural-networks-cnns-and-layer-types/>
10. Keiron O. and Ryan N. *An Introduction to Convolutional Neural Networks*. 2020. Prieiga per internetą: <https://arxiv.org/pdf/1511.08458.pdf>
11. *Simple Introduction to Convolutional Neural Networks*. 2019. Prieiga per internetą: <https://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac>
12. *Pooling Layer — Short and Simple*. 2020. Prieiga per internetą: <https://ai.plainenglish.io/pooling-layer-beginner-to-intermediate-fa0dbdce80eb>
13. *Transpose Convolution Explained for Up-Sampling Images*. 2020. Prieiga per internetą: <https://blog.paperspace.com/transpose-convolution/>
14. *Transposed Convolution Demystified*. 2020. Prieiga per internetą: <https://towardsdatascience.com/transposed-convolution-demystified-84ca81b4baba>
15. Jocher, G., Chaurasia, A., & Qiu, J. *Ultralytics YOLO (Version 8.0.0)*. 2023. Prieiga per internetą: <https://github.com/ultralytics/ultralytics>
16. *A Review of Popular Deep Learning Architectures: AlexNet, VGG16, and GoogleNet*. 2020. Prieiga per internetą: <https://blog.paperspace.com/popular-deep-learning-architectures-alexnet-vgg-googlenet/>

17. Jiangong N., Jiyue G.; Limiao D., Zhongzhi H.. *Monitoring the Change Process of Banana Freshness by GoogLeNet*, 2020. Prieiga per internetą: <https://ieeexplore.ieee.org/document/9296756>
18. *Hugging Face*. 2023. Prieiga per internetą: [https://huggingface.co/docs/transformers/model\\_doc/vit](https://huggingface.co/docs/transformers/model_doc/vit)
19. He, D., Zhao, H., Rudin, C., & Seltzer, M. *Towards Calibrated and Interpretable Neural Networks*, 2020. Prieiga per internetą: <https://arxiv.org/abs/2010.11929>
20. Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., ... & Le, Q. V. *Searching for MobileNetV3*. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019. Prieiga per internetą: [https://openaccess.thecvf.com/content\\_ICCV\\_2019/html/Howard\\_Searching\\_for\\_MobileNet\\_V3\\_ICCV\\_2019\\_paper.html](https://openaccess.thecvf.com/content_ICCV_2019/html/Howard_Searching_for_MobileNet_V3_ICCV_2019_paper.html)
21. Tan, M., Yu, A., Zhang, B., Cui, Y., & Le, Q. V. *EfficientNetV2: Smaller Models and Faster Training*. Proceedings of Machine Learning Research, 2021. 139, 10096-10110. Prieiga per internetą: <http://proceedings.mlr.press/v139/tan21a.html>
22. Oltean, M. *Fruits-360 dataset: new research directions*. 2021. Prieiga per internetą: [https://www.researchgate.net/profile/Mihai-Oltean-2/publication/354535752\\_Fruits-360\\_dataset\\_new\\_research\\_directions/links/6155f6b2a6fae644fbb4b4e1/Fruits-360-dataset-new-research-directions.pdf](https://www.researchgate.net/profile/Mihai-Oltean-2/publication/354535752_Fruits-360_dataset_new_research_directions/links/6155f6b2a6fae644fbb4b4e1/Fruits-360-dataset-new-research-directions.pdf)
23. Yang, J., Qin, S., & Lu, S. *Application of deep learning in intelligent transportation systems: A survey*. Data Science and Analytics, 3, 100019. 2021. Prieiga per internetą: <https://www.sciencedirect.com/science/article/pii/S2352340921009616>
24. Xiao, B., Nguyen, M. & Yan, W.Q. *Fruit ripeness identification using transformers*. Appl Intell 53, 22488–22499, 2023. Prieiga per internetą: <https://doi.org/10.1007/s10489-023-04799-8>
25. Nusrat S., Musfika J., Mohammad S. U. *An extensive dataset for successful recognition of fresh and rotten fruits*. Data in Brief, 2022. Prieiga per internetą: <https://doi.org/10.1016/j.dib.2022.108552>
26. Abayomi-Alli, O.O., Damaševičius, R., Misra, S. et al. *FruitQ: a new dataset of multiple fruit images for freshness evaluation*. Multimed Tools Appl 83, 11433–11460, 2024. <https://doi.org/10.1007/s11042-023-16058-6>
27. Alexander B., Vladimir I. I., Eugene K., Alex P., Mikhail D. and Alexandr A. K. *Albumentations: Fast and Flexible Image Augmentations*. Information, 2020. Prieiga per internetą: <https://www.mdpi.com/2078-2489/11/2/125>
28. TensorFlow. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. Prieiga per internetą: <https://www.tensorflow.org/>
29. CVAT.ai Corporation. *Computer Vision Annotation Tool (CVAT)*, 2023. Prieiga per internetą: <https://cvat.ai/>