

MAŠĪNU DZIĻĀS MĀCĪŠANĀS UN DATIZRACES PIELIETOŠANA AUGU UN PATOGĒNU MIJIEDARBĪBAS IZPĒTEI ĀBEĻU UN BUMBIERU KRAUPJA PATOSISTĒMĀS

APPLICATION OF DEEP LEARNING AND DATA MINING FOR THE STUDY OF PLANT-PATHOGEN INTERACTION: THE CASE OF APPLE AND PEAR SCAB

Gunārs Lācis¹, Inga Moročko-Bičevska¹, Olga Sokolova¹, Sergejs Kodors²

¹APP Dārzkopības institūts, ²Rēzeknes Tehnoloģiju akadēmija
gunars.lacis@llu.lv

Abstract. Fruit growing is an important niche in the structure of agriculture. Apple and pear are the most widely grown and economically significant fruit crops in the world and in Latvia, while the scab caused by *Venturia inaequalis* and *V. pyrina* are the most important diseases for these species. Considering environmental and food safety concerns, the high adaptability of pathogens and cost-effectiveness requirements, there is a need to change cultivation strategies by reducing the use of pesticides, promoting their precision and purposefulness. Smart or precision horticulture is a way to ensure this and provides a close linkage between research on local resources, environmental issues and information technologies, where common work can promote the development of fruit-growing. The aim of the study was the development of an integrated decision-making system using knowledge on plant-pathogen-environment interactions in apple/*V. inaequalis* and pear/*V. pyrina* pathosystems. The following objectives were defined to fulfil the aim: 1) application of semantic analysis and data mining for plant-pathogen interaction data in apple/*V. inaequalis* and pear/*V. pyrina* pathosystems; 2) development and implementation of an image-based deep learning system for early identification and evaluation of apple and pear scab; 3) development of IoT-system model for apple and pear monitoring. The results of this study provided the knowledge on plant-pathogen interaction mechanisms, their use for disease monitoring and prognosis.

Key words: *Malus*, *Pyrus*, *Venturia*, artificial intelligence, smart horticulture.

Ievads

Augļkopībai ir raksturīga augsta rentabilitāte un izaugsmes potenciāls, lai nodrošinātu tirgu ar daudzveidīgiem vietējiem pārtikas produktiem, izejvielām un inovatīviem nišas produktiem, un tā aizņem nozīmīgu vietu kopējā lauksaimniecības struktūrā. Ābeles (*Malus × domestica* Borkh.) un bumbieres (*Pyrus communis* L.) ir ekonomiski nozīmīgi augļaugi (Kaufmane et al., 2017). Ilgtspējīgu to audzēšanu būtiski ietekmē *Venturia inaequalis* (Cooke) G. Winter un *V. pyrina* Aderh. izraisītās slimības – ābeļu un bumbieru kraupis, kas rada nozīmīgus zaudējumus gan integrētajās, gan bioloģiskajās audzēšanas sistēmās (Belete, Boyraz, 2017; Sokolova et al., 2014).

Ābeļu un bumbieru kraupju izpētei ir sena vēsture. Ābelēm identificēta virkne rezistences gēnu, to donori, kas tiek izmantoti selekcijas programmās (Patocchi et al., 2020). Veikta daudzpusīga ābeļu kraupja patogēna/saimniekauga mijiedarbības izpēte, mazāk pētījumu veikti par bumbieru kraupi (Bouvier et al., 2012). Patogēnu populāciju mainības rezultātā selekcijā izveidotā rezistence tiek pārvarēta, kā arī lauka novērojumi uzrāda neatbilstības ar augu materiālā identificētajiem rezistences gēniem un novēroto izturību dabiskas inficēšanās apstākļos. Tās nevar precīzi izskaidrot, balstoties uz esošajiem datiem, tādējādi ierobežojot jaunu rezistences avotu atlasīšanu vai izveidi, sekmīgu un savlaicīgu selekcijas procesa nodrošināšanu. Līdzīgas problēmas novērotas arī bumbierēm, kam informācija par auga rezistenci un patogēnu ir ierobežotāka, rezistences selekcija notiek mazos apjomos, bez nozīmīgas ekonomiskās ietekmes. Lai gan augu rezistence tiek uzskatīta par vēlamāko slimību ierobežošanas metodi un mērķtiecīga, pret slimībām izturīgu šķirņu selekcija notiek daudzviet pasaulē, tomēr to īpatsvars dārzos nav liels, un joprojām plaši tiek izmantoti fungicīdi. Tas rada pretrunas starp rūpēm par vides piesārņojumu, iegūtās produkcijas drošību un augļkopības rentabilitāti, kā arī bažas par patogēnu augsto spēju veidot rezistenci pret lietotajiem fungicīdiem. Situācijās, kad nav iespējams pilnībā izvairīties no pesticīdu lietošanas, to izmantošanai jābūt pēc iespējas precīzākai un mērķtiecīgākai, ievērojami samazinot izmantošanas reižu skaitu un preparāta devu.

Viens no potenciālajiem risinājumiem ir viedo dārzkopības rīku izmantošana, kuru ieviešanā identificētas šādas problēmas:

- trūkst precīzu, zinātniski pamatotu fenotipēšanas paņēmieni, lai izstrādātu pietiekami precīzas slimību agrīnas identificēšanas un prognozēšanas sistēmas;
- trūkst arī bioloģiskās informācijas (patogēnu virulence, augu rezistence, augu un patogēnu mijiedarbības reakcijas) sasaistes ar sensoru datiem (dažādu spektru analīze: redzamā gaismā, IR, NIR, SWIR) to automatizētai analīzei;
- pieaugošs nestrukturētu datu apjoms – dažādos pētniecības projektos tiek iegūti dati par augu un patogēnu mijiedarbības aspektiem, taču netiek veikta to savstarpējo sakarību analīze, vides datu integrācija.

Šī iemesla dēļ pētījuma mērķis ir izstrādāt integrētu lēmumu pieņemšanas sistēmu, izmantojot zināšanas par augu, patogēnu un vides mijiedarbību ābeļu *V. inaequalis* un bumbieru *V. pyrina* patosistēmās.

Materiāli un metodes

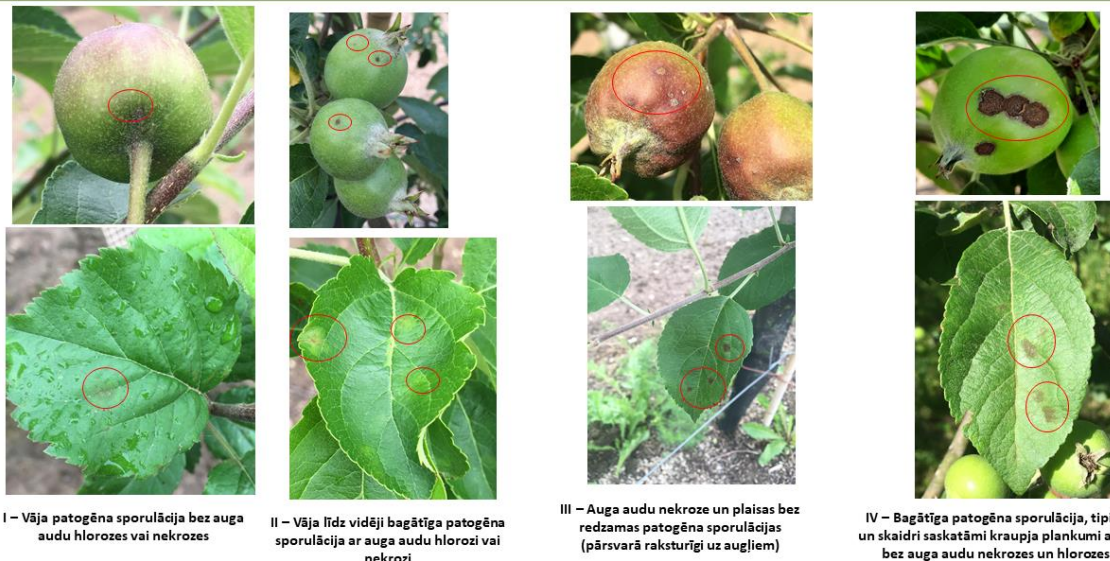
Viedās dārzkopības lēmumu pieņemšanas sistēmas izstrāde balstās uz 1) esošās pētnieciskās informācijas analīzi un 2) jaunu datu ieguvei, izmantojot dažādas attēlu ieguves tehnoloģijas.

Esošās pētnieciskās informācijas analīze veikta, izmantojot Dārzkopības institūtā pieejamos datus par dažādiem augu un patogēnu mijiedarbības aspektiem (augu rezistence dabiskas inficēšanās apstākļos, augu rezistences gēnu identificēšana, augu un patogēnu ģenētiskās daudzveidības analīze, patogēna virulences novērtēšana). Lai strukturētu un digitalizētu ekspertu zināšanas, izstrādātas semantiskās ontoloģijas, kas ļauj mākslīgajam intelektam veikt spriedumus par objektiem un to saitēm. Ontoloģiju izstrādē izmantoti *Protege* vai *WebProtege* rīki, zināšanas saglabājot *OWL 2* formātā.

Jaunu datu ieguvei veikta ābeļu un bumbieru kraupja simptomu fenotipēšana, izmantojot mobilās ierīces, iegūto attēlu analīze un slimības atpazīšana, izmantojot mašīnu dziļās mācīšanās metodi, kā arī hiperspektrālā attēlošana, agrīnai (pirms redzamiem simptomiem) infekcijas atpazīšanai. Attēlu iegūšana veikta lauka apstākļos (augļu dārzos), ābeļu veģetācijas sezonas laikā, izmantojot divu veidu mobilo sakaru ierīces ar dažādu kameras izšķirtspēju, izveidojot datu kopas ar attēliem, ar un bez infekcijas pazīmēm uz lapām un augļiem (skat. 1. att.). Savukārt agrīnās ābeļu kraupja infekcijas atpazīšanas metodes izstrāde veikta, izmantojot mākslīgo inficēšanu, kurā izmantoti podos augoši divgadīgi ābeļu stādi, kas inficēti ar *V. inaequalis* izolātiem. Inficēto lapu virsma dokumentēta ar *Specim IQ* (*Specim*, Somija) hiperspektrālo kameru, sākot ar trešo dienu pēc inokulācijas, kas turpināta katru dienu divu nedēļu garumā.

Venturia inaequalis– *Malus domestica* patosistēma: mijiedarbības pazīmes uz augļiem un lapām

I – agrīna slimības attīstības stadija saderīgas mijiedarbības rezultātā; II – agrīna slimības attīstības stadija nesaderīgas vai daļēji saderīgas mijiedarbības rezultātā; III – vēlna slimības attīstības stadija nesaderīgas vai daļēji saderīgas mijiedarbības rezultātā; IV – vēlna slimības attīstības stadija saderīgas vai daļēji saderīgas mijiedarbības rezultātā.



I – Vāja patogēna sporulācija bez auga audu hlorozes vai nekrozes

II – Vāja līdz vidēji bagātīga patogēna sporulācija ar auga audu hlorozi vai nekrozi

III – Auga audu nekroze un plaisas bez redzamas patogēna sporulācijas (pārsvarā raksturīgi uz augļiem)

IV – Bagātīga patogēna sporulācija, tipiski un skaidri saskatāmi kraupja plankumi ar vai bez auga audu nekrozes un hlorozes.

1. att. Attēlu iegūšanas metodika ābeļu kraupja simptomu atpazīšanai, izmantojot mašīnu dziļās mācīšanās metodi.

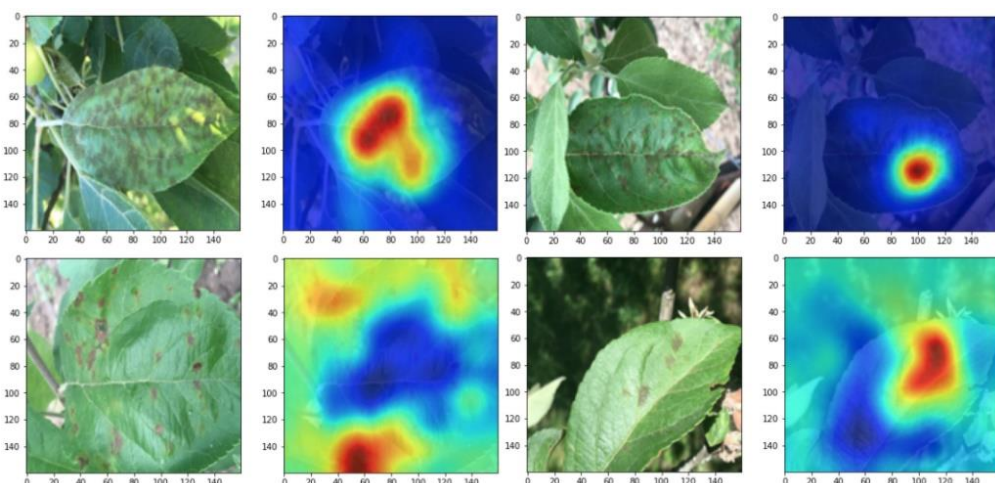
Fig. 1. Imaging methodology for recognizing apple scab symptoms using the machine deep learning method.

Rezultāti un diskusijas

Pētniecības aktivitātēs ir uzkrāti dati par dažādiem augu un patogēnu mijiedarbības aspektiem (augu rezistence dabiskas inficēšanās apstākļos, augu rezistences gēnu identificēšana, patogēna virulences novērtēšana, patogēna ģenētiskā daudzveidība), kuriem nav veikta to savstarpējās saistības analīze, kā arī nav iesaistīta vides un meteoroloģisko datu izmantošana. Paredzams, ka integrētas šādu datu sistēmas izveide, to semantiskā analīze, īstenojot datizraces metodes, ļaus identificēt jaunas likumsakarības augu, patogēnu un vides mijiedarbībā, kā arī nodrošinās zināšanu bāzi slimību kontroles lēmumu pieņemšanai. Pētījuma laikā izstrādāta ontoloģija, kas iekļauj informāciju par patogēniem, šķirnēm, dārzu ģeotelpiskajām īpašībām un laika apstākļiem. Šī ontoloģija var tikt izmantota viedā dārza izstrādē, lai ar lietu interneta palīdzību iegūtu datus un izmantotu tos prognozēšanā vai lēmumu pieņemšanas procesā. Savukārt, lietojot semantiskos vaicājumus, var realizēt elektronisko ekspertu, kas būs spējīgs atbildēt uz lietotāja jautājumiem vai īstenot loģiskus spriedumus.

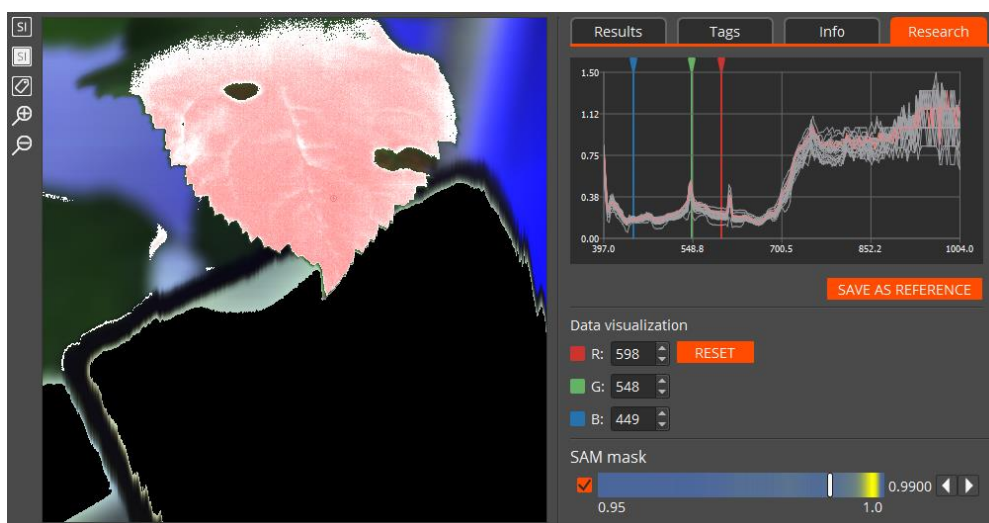
Esošo zināšanu pilnveidošanai pētījumā veikta ābeļu un bumbieru kraupja simptomu fenotipēšana, izmantojot mobilās ierīces, iegūto attēlu analīzi un slimības atpazīšanu, izmantojot mākslīgo intelektu, kā arī hiperspektrālo attēlošanu agrīnai infekcijas atpazīšanai un digitālās ekspertu sistēmas izstrādi, balstoties uz semantisko risinājumu un domēna zināšanu ontoloģijas lietošanu.

Viedais risinājums izstrādāts vairākos posmos. Sākotnēji analizēti esoši pētījumi, pieejamās attēlu datu kopas un mākslīgajā intelektā balstīti risinājumi. Analīzes rezultātā identificēti augu slimību atpazīšanas risinājumi, kas izstrādāti, izmantojot *PlantVillage* (Hughes, Salathé, 2015) un *Plant Pathology 2020 (FGVC7)* (Thapa et al., 2020). Tomēr šīm abām esošajām datu kopām piemīt savi trūkumi – abas datu kopas satur tikai inficētu lapu attēlus, bet nesatur attēlus ar augļiem to agrīnā attīstības stadijā. Savukārt *PlantVillage* attēli nav uzņemti dabiskos apstākļos, tāpēc šajā pētījumā bija nepieciešams izstrādāt paraugu fotografēšanas metodoloģiju un iegūt domēna datus mašīnāpmācībai. Lai veiktu mākslīgā intelekta apmācību, ābeļu kraupja simptomu atpazīšanai izveidota datu kopa, kurā iekļauti attēli ar un bez infekcijas pazīmēm uz lapām un augļiem. Līdztekus analizētas esošas konvolūcijas neironu tīklu arhitektūras, kas paredzētas mobilajām iekārtām (Kodors et al., 2020). Pētījums papildināts ar pārnestās apmācības (*transfer learning*) risinājumu eksperimentālu analīzi, jo datu savākšanas gadā novērojumu veikšanas vietā ābeļu kraupja attīstības pakāpe bija nepietiekama, kas ietekmēja plānotās datu kopas izmēru. Veicot tehnoloģiju izvēli, iespējamie varianti salīdzināti, izmantojot statistiskas analīžu *Mann–Whitney–Wilcoxon* testu. Savukārt dziļā mašīnāpmācīšanās (*deep learning*) īstenoja, pielietojot *Keras* satvaru un *CUDA* atbalstītas videokartes. Neironu tīklu precizitāte un atpazīšanas kvalitāte pārbaudīta, salīdzinot vairākus parametrus – atpazīšanas precizitāti (*training accuracy*), validācijas precizitāti (*validation accuracy*), Kohena kapa koeficients (*Cohen's Kappa*), uzmanības un aizsprostu kartes (*saliency and occlusion maps*) (skat. 2. att.).



2. att. Kvalitātes kontrole: attēlu vietas ar paaugstinātu neironu tīklu uzmanību.
 Fig. 2. Quality control: image areas with increased neural network focus.

Pētījumā laikā izveidotas divas attēlu datu kopas: lauka un dabiskās infekcijas apstākļos iegūto digitālo attēlu datu kopa (1181 attēls) mākslīgā intelekta apmācībai, kā arī laboratorijas apstākļos un mākslīgās infekcijas rezultātā iegūtā hiperspektrālo datu kopa (300 attēli) agrīnās kraupja identifikācijas metodes izstrādei (skat. 3. att.). Kraupja simptomu atpazīšanai pētījumā laikā salīdzinātas šādas konvolūcijas neironu tīklu arhitektūras: *AlexNet*, *AlexNet* modifikācija, aizvietojo *Conv* slāņus ar *SeparateConv* variantu, *MobileNetV1* un *MobileNetV2*. Rezultātā konstatēts, ka *MobileNetV2* arhitektūra ir vispiemērotākā gan pēc atmiņas, gan pēc apstrādes ātruma un atpazīšanas precizitātes. Ņemot vērā, ka dārzu apsekošanas laikā izdevās savākt nelielu apmācības attēlu datu kopu (208 attēli ar inficētiem āboliem un 973 – ar inficētām ābeļu lapām), salīdzinātas brīvi pieejamas datu kopas pārnestās apmācības īstenošanai. Eksperimenta laikā salīdzinātas piecas datu kopas: *CIFAR-10*, *CIFAR-100* (Krizhevsky, 2009), *PlantVillage* (Hughes, Salathé, 2015), *Fruits360* (Mureşan, Oltean, 2018) un *iFood251X* (Kaur et al., 2019). Attālums tika mērīts, izmantojot *Earth Mover's Distance* saskaņā ar Cui u. c. aprakstītu metodoloģiju (Cui et al., 2018). Vistuvākais variants pētījumā iegūtajai datu kopai bija *iFood251X*. Jāuzsver, ka eksperimenta rezultāti sakrīt ar Y. Cui un līdzautoru novērojumiem, ka pārāpmācītu neironu tīklu attālums ietekmē neironu tīkla precizitāti.



3. att. Hiperspektrālā ābeļu kraupja simptomu identifikācija un spektrālā analīze agrīnai infekcijas atpazīšanai.

Fig. 3. Identification and spectral analysis of apple scab symptoms for early detection of infection by hyperspectral data acquisition.

Turpmākajos eksperimentos tiek plānots apvienot izvēlētās tehnoloģijas vienā apmācības metodē. Ievērojot, ka 2020. gadā tika publicēta *Plant Pathology 2020 (FGVC7)* domēna datu kopa, to apvienošana ar iegūtajiem datiem ir labs pamats *TRL6* prototipa izstrādei un turpmākai validēšanai, izmantojot iedzīvotāju iesaistes (*crowdsourcing*) pieeju.

Secinājumi

1. Datizraces metodes ir piemērotākais risinājums likumsakarību identificēšanai augu, patogēnu un vides mijiedarbībā un zināšanu bāzes nodrošināšanai slimību kontroles lēmumu pieņemšanā.
2. Ir būtiski pilnveidot fenotipēšanas metodes, lai izstrādātu agrīnas un pietiekami precīzas ābeļu kraupja identificēšanas un prognozēšanas sistēmas, kā arī lai sasaistītu šo informāciju ar esošo augu un patogēnu mijiedarbības bioloģisko informāciju tās automatizētai analīzei.

Pateicība

Pētījums tiek īstenots projekta Nr. LZP-2019/1-0094 "Mašīnu dziļās mācīšanās un datizraces pielietošana augu un patogēnu mijiedarbības izpētei: ābeļu un bumbieru kraupja patosistēmas" ietvaros.

Izmantotā literatūra

1. Belete T., Boyraz N. (2017) Critical Review on Apple Scab (*Venturia inaequalis*) Biology, Epidemiology, Economic Importance, Management and Defense Mechanisms to the Causal Agent. *J. Plant Physiol. Pathol.*, Vol. 5, p. 2.
2. Bouvier L., Bourcy M., Boulay M., Tellier M., Guérif P., Denancé C., Durel C. E., Lespinasse Y. (2012). A New Pear Scab Resistance Gene *Rvp1* from the European Pear Cultivar 'Navara' Maps in a Genomic Region Syntenic to an Apple Scab Resistance Gene Cluster on Linkage Group 2. *Tree Genetics & Genomes*, Vol. 8, p. 53–60.
3. Cui Y., Song Y., Sun C. Howard A., Belongie S. (2018). Large Scale Fine-Grained Categorization and Domain-Specific Transfer Learning. **In:** *Proceedings of International Conference: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Salt Lake City, USA, p. 4109–4118.
4. Hughes D. P., Salathé M. (2015) An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics. [Tiešsaiste] [skatīts 2021. g. 16. jūn.] Pieejams: <https://arxiv.org/abs/1511.08060>.
5. Kaufmane E., Skrīvele M., Ikase L. (2017) Fruit Growing in Latvia – Industry and Science. *Proc. of the Latvian Academy of Sciences*, Vol. 71, p. 237–247.
6. Kaur P., Sikka K., Wang W., Belongie S., Divakaran A. (2019) FoodX-251: A Dataset for Fine-grained Food Classification. [Tiešsaiste] [skatīts 2021. g. 16. jūn.] Pieejams: <https://arxiv.org/abs/1907.06167>.
7. Kodors S., Lacis G., Zhukov V., Bartulsons T. (2020) Pear and Apple Recognition Using Deep Learning And Mobile. **In:** *Proceedings of 19th International Scientific Conference Engineering for Rural Development*. Latvia University of Life Sciences and Technologies, Jelgava, Latvia, p. 1795–1800.
8. Krizhevsky A. (2009) Learning Multiple Layers of Features from Tiny Images. [Tiešsaiste] [skatīts: 2021. g. 11. febr.]. Pieejams: <https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf>.
9. Mureşan H., Oltean M. (2018) Fruit Recognition from Images Using Deep Learning. *Acta Universitatis Sapientiae, Informatica*, Vol. 10(1), p. 26–42.
10. Patocchi A., Wehrli A., Dubuis P-H., et al. (2020) Ten Years of VINQUEST: First Insight for Breeding New Apple Cultivars With Durable Apple Scab Resistance. *Plant Disease*, Vol. 104, p. 2074–2081.
11. Sokolova O., Moročko-Bičevska I., Bankina B. (2014) Review of the Pear Scab Caused by *Venturia pyrina*. **In:** *Annual 20th International Scientific Conference Proceedings, "Research for Rural Development"* Jelgava, Latvia, 21–23 May 2014. Vol. 1, p. 26–33.
12. Thapa R., Zhang K., Snaveley N., Belongie S., Khan A. (2020) The Plant Pathology Challenge 2020 Data Set to Classify Foliar Disease of Apples. [Tiešsaiste] [skatīts 2021. g. 16. jūn.] Pieejams: <https://arxiv.org/abs/2004.11958>.