

MASHINALI O'QITISH MODELLARI UCHUN TASVIR DESKRIPTORLARI ASOSIDA YUQORI ANIQLIKDAGI O'QUV TANLANMALARINI SHAKLLANTIRISH ALGORITMLARI

Madraximov A.X., Murayeva H.M.

Muhammad al-Xorazmiy nomidagi TATU, dotsent

Muhammad al-Xorazmiy nomidagi TATU, assistant

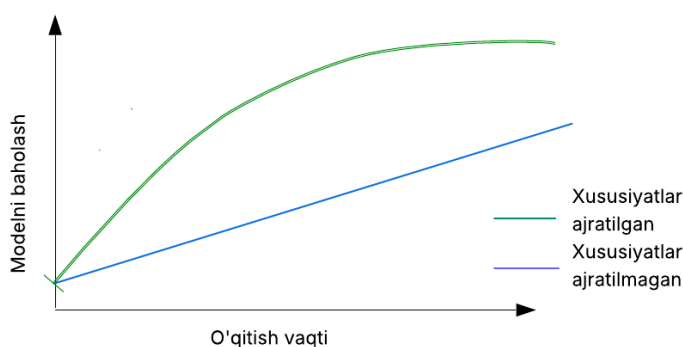
ANNOTATSIYA: Mazkur tadqiqot ishi timsollarni tanib olish sohasida obyektning harakterlovchi xususiyatlarni ajratish algoritmilarini takomillashtirishga qaratilgan bo'lib, mazkur ishda Kaggle platformasidagi "Pomidor kasalliklari" (PlantVillage) ma'lumotlar to'plamidan 10 000 ta rasm ishlatilgan bo'lib, ulardan 29 ta ma'lumot beruvchi xususiyatlardan iborat o'quv to'plami shakllantirilgan. Bu yondashuv issiqxona va ochiq maydonlarda ekinlarni erta tashxislashda va himoya choralarini prognozlashda muhim amaliy ahamiyatga ega. Umotlar to'plamidan 10 000 ta rasm ishlatilgan bo'lib, ulardan 29 ta ma'lumot beruvchi xususiyatlardan iborat o'quv to'plami shakllantirilgan. Natijalar shuni ko'rsatadiki, taklif qilingan xususiyatlarni saralash usuli modelning aniqlik koeffitsientini oshirish bilan birga hisoblash vaqtini sezilarli darajada tejaydi. Bu yondashuv issiqxona va ochiq maydonlarda ekinlarni erta tashxislashda va himoya choralarini prognozlashda muhim amaliy ahamiyatga ega.

Kalit so'zlar: tasvirlarga ishlov berish, xususiyatlarni ajratish, pomidor kasalliklari, mashinali o'qitish, rang deskriptorlari, tekstura tahlili.

O'simlik kasalliklarining tarqalish dinamikasi va ularning hosil sifatiga ta'sirini aniq baholash seleksiya, o'simliklar himoyasi va fungitsidlar samaradorligini tahlil qilishda fundamental ahamiyatga ega. Kasallik o'choqlarini o'z vaqtida va ishonchli aniqlash nafaqat issiqxona va ochiq maydonlarda himoya choralarini optimallashtirish, balki patologik jarayonlarning vaqt va makon bo'yicha rivojlanishini modellashtirish (prognozlash) imkonini beradi.

O'simlik kasalliklarini vizual ma'lumotlar asosida diagnostika qilishda esa tasvirlardan yuqori informativlikka ega xususiyatlarni ajratish va optimallashtirish hal qiluvchi ahamiyatga ega. Rang, tekstura va geometrik deskriptorlarning to'g'ri tanlanishi modelning aniqlik koeffitsientini oshirish bilan birga, hisoblash resurslarini tejash imkonini beradi. Shu bois, tasvirlarga raqamli ishlov berish jarayonida tayanch xususiyatlarni saralash algoritmilarini takomillashtirish bugungi kunda mashinali o'qitish sohasidagi eng muhim vazifalardan biri hisoblanadi.

Xususiyatlar-bu tasvirdagi obyektlarni yagona belgilar bilan belgilashga yordam beradigan atributlar. Ular umuman olganda, to'liq tasvirlar yoki tasvir to'plamlarini farqlashda muhim ahamiyatga ega bo'ladi.



1-rasm. Modelni baholash va o'qitish vaqti

Tasvirlardan xususiyatlarni ajratib olishning asosiy usullari quyidagilardan tashkil topgan:

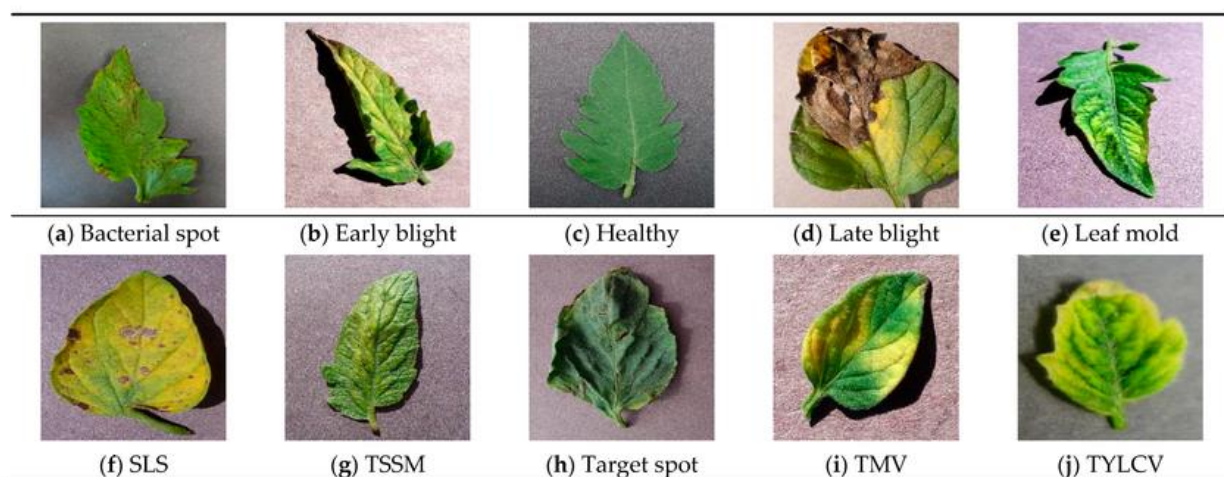
- Rang;
- Tekstura;
- Shakl;
- Gradient;
- Statistik;
- Transformatsiya;
- Chuqur o'qitish.

1-jadval. Obyektni harakterlovchi belgilarni ajratishning analitik ifodalari

1.	<i>Color moments:</i> Mean. Represents the average color value in the image. $M_t(s_j) = \sum_{u=1}^N \frac{1}{N} I_{t,u},$ Where I_{tu}- the intensity values of pixel u of t-th color channel. N- the number of pixels.	<i>Red-Green Ratio, Blue-Green Ratio, Red-Blue Ratio</i> $RGR_j = \frac{1}{N} \sum_{u=1}^N \frac{I_{R,u}}{I_{G,u} + \epsilon}$ $BGR_j = \frac{1}{N} \sum_{u=1}^N \frac{I_{B,u}}{I_{G,u} + \epsilon}$ $RBR_j = \frac{1}{N} \sum_{u=1}^N \frac{I_{R,u}}{I_{B,u} + \epsilon},$
2.	Standard deviation. Represents the square root of the color distribution variance $SD_t = \sqrt{\frac{1}{N} \sum_{u=1}^N (I_{t,u} - M_t)^2}$	GLCM features (texture) <i>Contrast</i> $f_i(s_j) = \frac{1}{3} \sum_{t=1}^3 \sum_{I_{t,u}=0}^{L-1} \sum_{I_{t,v}=0}^{L-1} (I_{t,u} - I_{t,v})^2 P_{j,t}^\theta(I_{t,u}, I_{t,v})$
3.	Skewness. The skewness level of color distribution: $SK_t = \sqrt[3]{\frac{1}{N} \sum_{u=1}^N (I_{t,u} - M_t)^3}$	<i>Homogeneity</i> $f_i(s_j) = \frac{1}{3} \sum_{t=1}^3 \sum_{I_{t,u}=0}^{L-1} \sum_{I_{t,v}=0}^{L-1} \frac{P_{j,t}(I_{t,u}, I_{t,v})}{1 + (I_{t,u} - I_{t,v})^2}$
4.	Kurtosis. The shape of color distribution: $K_t = \sqrt[4]{\frac{1}{N} \sum_{u=1}^N (I_{t,u} - M_t)^4}$	<i>Energy</i> $f_i(s_j) = \frac{1}{3} \sum_{t=1}^3 \sqrt{\sum_{I_{t,u}=0}^{L-1} \sum_{I_{t,v}=0}^{L-1} P_{j,t}(I_{t,u}, I_{t,v})^2}$
5.	<i>Excess Green Index (ExG)</i> $ExG_j = \frac{1}{N} \sum_{u=1}^N (2I_{G,u} - I_{R,u} - I_{B,u})$	<i>Correlation</i>

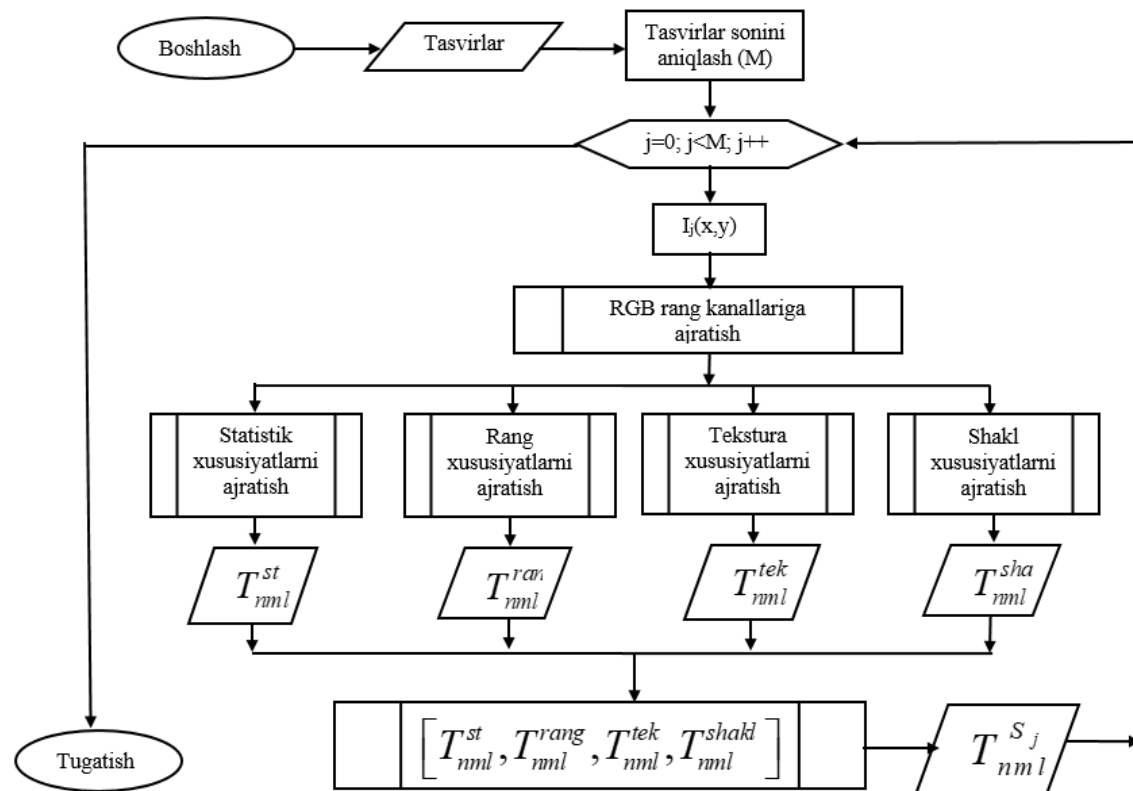
	Where ExG_j – For the j-th object (image), it represents the Excess Green Index (ExG) value $I_{G,u}; I_{R,u}; I_{B,u}$ – The intensity values of the u-th pixel in the green, red, and blue channels, respectively.	$f_i(s_j) = \frac{1}{3} \sum_{t=1}^3 \frac{\sum_{I_{t,u}=0}^{L-1} \sum_{I_{t,v}=0}^{L-1} (I_{t,u} - \mu_{t,u})(I_{t,v} - \mu_{t,v}) P_{j,t}(I_{t,u}, I_{t,v})}{\sigma_{t,u} \sigma_{t,v}}$
6	<i>Green Leaf Index (GLI)</i> $GLI_j = \frac{1}{N} \sum_{u=1}^N \frac{(2I_{G,u} - I_{R,u} - I_{B,u})}{(2I_{G,u} + I_{R,u} + I_{B,u})}$	

Obyektlarni harakterlovchi xususiyatlarni ajratishga misol tariqasida Kaggle.com ochiq platformasidan PlantVillage ma'lumotlar ba'zasidan "Tomato diseases" ma'lumotlar ba'zasi namuna sifatida olindi. Ushbu ma'lumotlar ba'zasida turli kasalliklardan iborat 10 ta sinf mavjud bo'lib, jami 10000 ta tasvirdan tashkil topgan. Ular mos ravishda 9 xil kasallangan barg tasvirlaridan tashkil topgan sinflar hamda 1 ta sog'lom barg tasvirlaridan iborat sinf. Mazkur ma'lumotlar ba'zasi quyidagi kasalliklardan iborat:



2-rasm. "Pomidor kasalliklari" datasetidan sinf namunalari

Yuqoridagi analitik ifodalardan foydalanib quyidagicha ansambl algoritmi shakllantirildi:



2-rasm. Obyektni harakterlovchi xususiyatlarni ajratish algoritmi

Yuqoridagi algoritimni qo'llash natijasida quyidagicha o'quv tanlanma shakllantirildi:

Belgilar Obyektlar	x_1 (Mean_R)	x_2 (Mean_G)	x_3 (Mean_B)	x_4 (Std_R)	x_{26} (HOG 1)	x_{27} (HOG 2)	x_{28} (HOG 3)	x_{29} (HOG 4)	Sinflar
S_1	0.59361	0.61396	0.64023	0.5394 3	0.3491 4	0.0681 8	0.3197 8	0.4083 6	Tomato__Leaf_M old
S_2	0.6447	0.57222	0.54721	0.2326 5	0.3964 7	0.2747 1	0.3263 3	0.2586	Tomato__Leaf_M old
S_3	0.55794	0.57689	0.58588	0.4362 8	0.2706 6	0.0886	0.4056 8	0.1764 2	Tomato__Leaf_M old
S_4	0.61575	0.56545	0.52266	0.3726 8	0.2922 8	0.2251 8	0.3012 7	0.0989 9	Tomato__Leaf_M old
S_5	0.67492	0.64564	0.62378	0.6164	0.4327 7	0.2977 8	0.0866 4	0.1324 7	Tomato__Leaf_M old
S_6	0.71788	0.56114	0.48643	0.5116 9	0.2809 3	0.1390 7	0.3903 9	0.2901 8	Tomato__Leaf_M old

S_7	0.63385	0.63622	0.65761	0.56107	...	0.31839	0.12527	0.25932	0.31149	Tomato__Leaf_Mold
S_8	0.64316	0.63685	0.66585	0.50153	...	0.17383	0.25181	0.59311	0.22089	Tomato__Leaf_Mold
...	...	...	...	...	...	...	...	...	...	...
S_{10015}	0.06469	0.05292	0.18426	0.27357	...	0.27652	0.15344	0.35713	0.24154	Tomato__Late_blight
S_{10016}	0.13266	0.10435	0.22091	0.22296	...	0.37853	0.13928	0.41065	0.23373	Tomato__Late_blight
S_{10017}	0.07349	0.04293	0.22059	0.28286	...	0.28575	0.29701	0.21098	0.29279	Tomato__Late_blight
S_{10018}	0.15399	0.13593	0.14107	0.28838	...	0.41162	0.04924	0.14126	0.24276	Tomato__Late_blight

Ushbu ishda raqamli tasvirlarni qayta ishlash orqali o'simlik tasvirlaridan rang, tekstura va shakl tavsiflovchilarini ajratib olish algoritmlari takomillashtirildi. "PlantVillage" ma'lumotlar to'plamidagi 10 000 ta pomidor barglari tasvirlari bo'yicha o'tkazilgan testlar ma'lumot beruvchi xususiyatlarni saralash tizimning barqaror ishlashini ta'minlashini ko'rsatdi. Rivojlangan ansambl algoritmi va shakllantirilgan o'quv tanlovi kasallik tarqalishining dinamikasini bashorat qilishda yuqori samaradorlikka ega, shu bilan birga mashinani o'rganish modellarining aniqligini oshiradi.

FOYDALANILGAN ADABIYOTLAR

1. Ali H, LaliMI, NawazMZ, SharifM, Saleem BA (2017) Symptom based automated detection of citrus diseases using color histogram and textural descriptors. Comput Electron Agric 138:92– 104. <https://doi.org/10.1016/j.compag.2017.04.008>
2. Deng Y, Yu Y (2019) Self-feedback image retrieval algorithm based on annular color moments. EURASIP J Image Video Process 2019(1):1–13. <https://doi.org/10.1186/s13640-018-0400-9>
3. Souza L, Ebigbo A, Probst A, Messmann H, Papa JP, Mendel R, Palm C (2018) Barrett's Esophagus identification using color co-occurrence matrices. In: 2018 31st SIBGRAPI conference on graphics, patterns and images (SIBGRAPI), pp 166–173. <https://doi.org/10.1109/sibgrapi.2018.00028>
4. Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L. R., Davis, C. E., and Dandekar, A. M. 2014. Advanced methods for plant disease detection. A review. **Agron. Sustain. Dev.** 35:1-25. <https://doi.org/10.1007/s13593-014-0246-1> [Crossref](#)[Web of Science](#)[Google Scholar](#)
5. Suhang Wang, Jiliang Tang, and Huan Liu, Feature selection, USA, 2017.