INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 04, April 2024, pp: 2582-2590

e-ISSN : 2583-1062

Impact Factor: 5.725

PEST DETECTION AND CLASSIFICATION IN PEANUT CROPS USING CNN, MFO, AND EVITA ALGORITHMS

Sunita B[1], Yathin Kumar M[2], Naveen V[3], Sharath TR[4], Vikas MG[5]

[1]Head of Dept of CSE, Sambhram Institute of Technology, Bengaluru Urban, Karnataka, India.

[2,3,4,5] Final Year Students, Dept. Of CSE, Sambhram Institute of Technology, Bengaluru Urban, Karnataka, India.

DOI: https://www.doi.org/10.58257/IJPREMS33750

ABSTRACT

The growth of vision transformer (ViT) methods have been quite enormous since its featuresprovide efficient outcome in image classification, and identification. Inspired of this beneficial, this paper propose an Enhanced vision transformer architecture (EViTA) model for pest identification, segmentation,and classification. The as of late found that, compare to machine learning, Convolutional neural network algorithms the ViT has providing trusted results on image classification. Motivated by this, in this paper, we concentrate on the best way to learn dual barch segment representations in ViT models for image arrangement. Based upon its features, here propose a double layer transformer encoder to integrate pest image segments of various sizes of pest images to create more grounded image highlights. The current study uses, three pest datasets that affects peanut crops such as Aphids (IP102 Dataset), Wireworm (IP102 Dataset), and Gram Caterpillar collected from public available repository. Our methodology processes small segmentand huge segment of tokens with two separate parts of various computational intricacy also, these tokens are then combined simply by consideration numerous times to complete one another. The taken datasets' are preprocessed utilizing the characteristic by using moth flame optimization (MFO), and flatten the images by using linear projector methodology to enhance the missing quality in pest images, and afterwardnormalization methods are executed to switch it over completely in to mathematical arrangement. This processed information is standardized further utilizing the self attention in StandardScaler procedures are carried out for choosing the ideal highlights in the dataset accordingly having huge effect towards affectingpest image predictions. These ideal highlights are at last taken care of into the EViTA model and the outcomes created are considered in contrast to the cutting edge models which at last legitimize the predominance of the proposed EViTA PCA MFO model in pest image prediction with high accuracy rate. Broad trials show that our methodology performs better compared to or on standard with a few simultaneous deals with visiontransformer, notwithstanding productive CNN models.

1. INTRODUCTION

Agriculture is critical to feeding both human and livestock populations worldwide. Agriculture's role in clean energy generation has expanded with the adoption of environmen- tally friendly artificial intelligence (AI), and Internet of Things (IoT) technologies. Moreover, farming is additionallythe well-spring of natural substances utilized in making mate- rials, synthetic compounds and drugs. Regardless of a minor 15% expansion in how much land under horticultural use between the 1960s and the early piece of 100 years, farmingcreation tripled. This was ascribed to reception of pesticidesand fertilizers, as well as precision farming and the devel- opment of higher yielding crop and livestock varieties [1]. As of late, the rate of development in rural creation has beendeclining [2]. This pattern, combined with arising difficultieslike environment change, populace development [3], countryto metropolitan relocation.

Plant diseases and insect pests have always been one of the main factors restrict ing the sustainable development of agriculture. On the one hand, plant diseases have caused a lot of economic losses and even caused famine. According to the United Nations Food and Agriculture Organization (FAO) estimates, 10% of cereal production is lost due to diseases, and 12% of cotton production is lost due to diseases perennially. Accordingly, the economic loss caused by pests in the world is as high as 120 billion dollars every year, equivalent to China's agri cultural output value, 1/3 of the United States, two times that of Japan, and more than four times that of the United Kingdom [1]. Production cuts due to epidemics have been a global problem. (1) In 1970, the epidemic of corn blight in the United States caused a loss of 1 billion US dollars [2]. (2) In 1990, China's wheat stripe rust epidemic reduced rice production by 2.5 billion kilograms. In 1993, the rice blast epidemic in Chinese southern rice areas reduced rice production by 15 billion kilograms [3]. (3) In 1845, the Irish famine that shocked the world was caused by the potato late blight epidemic. In 1942, a large rice area in Bangladesh suffered from flax spot disease, and by 1943, 2 million people died of starvation [4].

(4) Rice dwarf disease was prevalent in some areas of Japan at the end of the 19th century, and more than 10,000

people starved to death due to it [5]. (5) Cocoa swollen branch disease is extremely devastating in Africa. Ghana alone has cut down 179 million diseased trees from 1946 to 1981 [6]. On the other hand, pests are another factor restricting the

development of the agricultural economy. There are more than 400 kinds of pests that have been found in China, and more than 40 kinds are more common in the northern region. It is divided into several major categories, such as leaf-eating, under ground, gnawing, and sucking pests [7]. Compared with plant diseases, pests cause damage to more parts of plants, which will affect the roots, stems, leaves and fruits of plants. According to relevant statistics and analysis, from 2006 to 2015, the affected area of crop diseases and insect pests in China ranged from 463 million hm2 to 507.5 million hm2. According to the forecast of the National Agricultural Technology Extension Service Center, in 2020, the affected area of crop diseases and insect pests in China will reach 300 million hm2, which will cause colossal food and economic losses [8]

Peanut is a harvest with different purposes and an excep- tionally rich dietary benefit. As the primary oil also, financial yield of our Government, its developing region is contin- ually expanding, and it has turned into the second biggest developed crop in India [5]. The actual nut also, the multi- faceted field environments simplify the leaves to be polluted by microorganisms. The microorganism can spread rapidly through ordinary factors and has a high regenerative cutoff. The essential variable impacting its proliferation is the moist-ness of peanuts during the seedling stage [6]. Ailments of nut leaves can decrease the yield and nature of peanuts by annihi-lating the green tissue and chlorophyll in the leaves [7]. Fake unmistakable verification of nut leaf contaminations requirescapable data, and it is easy to misdiagnose them basically by counterfeit visual insight. Thusly, nut sicknesses can't bedissected and treated in time. The method for controlling nut disease is to break down the sickness type quickly and exactly, and subsequently take relating control measures in time [8].

There are various techniques for distinguishing plant sick- nesses in their beginning phases. The conventional techniquefor plant sickness discovery is unaided eye observing, whichis insufficient what's more, erroneous for huge crops [9]. Accordingly, illness identification in leaves is a significant subject that gives many advantages in checking enormous fields of harvests. Peanut sickness can influence yield and quality by harming the green layer from the leaves. The method for controlling these peanut infections is to quickly and unequivocally identify the sickness type and afterward carry out suitable restorative activities in a convenient man-ner [10]. Utilizing the beneficial of advanced DNN, CNN, ML, and vision transformer (ViT) algorithms, the recogni- tion of sickness is productive, consumes less time, and is exact [10].

The current study focuses on predicting the peanut disease in real time environment. Even Machine learning, and CNN algorithms are given suitable for image classification,segmentation, and identification in effective manner, the accuracy which it has producing is not up the level [11].In order to achieve an efficient result in this work proposed an Enhanced Vision Transformer Architecture (EViTA) for analyzing, classifying, segmenting, and identifying peanut pest based on images. In EViTA method, the gathered inputimages are segmented into number of segments for easy processing. Those processed images were embedded into positional encoding, and feed into distinct transformer layersfor accurate identification of nature of peanut pest.

Experiments were conducted in this study using the pub- licly available Aphids, Wireworm, and Gram Caterpillar pestdatasets that affects peanut crops in maximum. To extract the features from the chosen datasets the MFO algorithm is usingin this work. The extracted features are put into the block called Extra arrangement segment (EAS), the dataset contains the best features that would have significant effect peanut growth. Finally, EViTA method is fed with the extracted data with the objective to predict affected peanut crops, and also help to increase the growth of peanut crops. The significant commitments of the proposed model are as follows,

1) CNN is used to predict the pest infection in peanut crops.
2) In order to increase the optimum prediction rate, MFOis used to select the most appropriate features.
3) MFO and state-of-the-art techniques are thoroughly examined.
4) Experiments show that the proposed model dominatesother popular EViTA methods, signifying the beneficial effect of MFO integration with the EViTA methods.

The structure of the paper includes the following. Section IIexamines the survey of writing. Section III sets out the starterwork and the strategy proposed. Section IV explores the picked dataset with proposed techniques results and conver-sations, and Section V depicts the conclusion and suggestions for future work.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 04, April 2024, pp: 2582-2590

e-ISSN : 2583-1062

Impact Factor: 5.725

2. METHODOLOGY

A. CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN is right now perhaps of the most well known model andhas displayed their extraordinary execution on many image grouping issues in prominent field [30]. The idea of sharing loads in CNN makes a powerful image segment by finding powerful elements in the images and lessen the dissipating tendency issue. The designing of a normal CNN is given in Fig. 2. The development of CNN integrates convolution layer, pooling layer, and totally related layer. The convolutional layer goes about as channels moreover, the chief errand is to remove features from the insect images. The convo- lutional layer is followed by pooling layer, which performs down looking at and holds the fundamental information in theThe rundown of flames is displayed in wherein in a state of harmony with the past conditions x signify the quantity of moths and n shows the amount of angles mirrors the group for taking care of the flame wellness values.insect images. This layer reduces the spatial size of depiction as well as the amount of limits and prevents overfitting which makes the model more capable. The last layer is the totally related layers that use a ReLu establishment capacity and takes the certain level components from the insect images for gathering them into various classifications with marks .

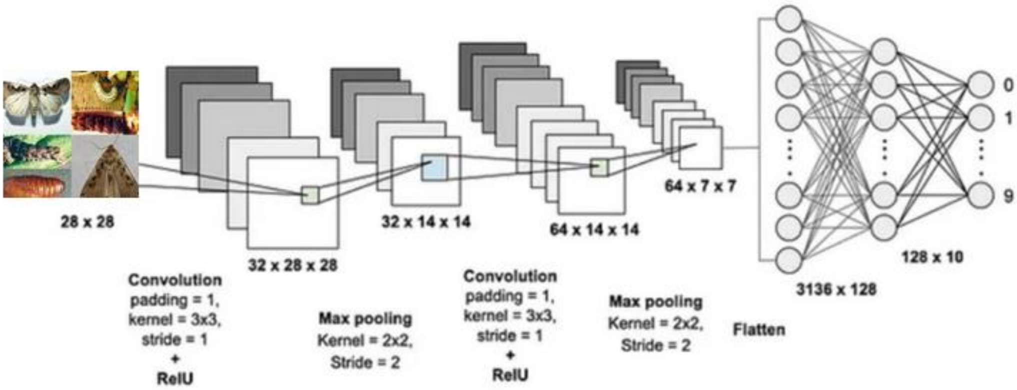

FIGURE 2. Pest classification in CNN.

B. MOTH FLAME OPTIMIZATION (MFO) ALGORITHM

There exists pretty much 160,000 particular kinds of bug species in the world. Amidst nature, Moths are bugs having a spot with the butterfly family species. A presence of a moth twirls around fulfillment of two principal targets – hatchlings and adults. The hatchlings gets exchanged over totally to a moth inside the housings. Moths investigate including a very specific yet useful part at night. Moths utilize dusk to fly in the late night utilizing the cross bearing framework that assists them with investigating. The moths fly keeping a specific point with the moon which helps them with exploring longer distance in a straight manner. Yet the cross heading part is seen as extremely convincing, a moth flying development is often seen to be unpredictable and twisting around the light. These exclusions occur inside seeing fake lights, wherein moths see the phony light and cross straight towards such light keeping a specific point. However, the fake light being in such nearness confounds the moths to participate in a lethal winding fly prompting destructive results.

The MFO algorithm is following meta heuristic calculation that follows comparative personal conduct standard as moths and subsequently assumes a critical part in highlight determination [19]. The ideal place of the moths is introduced utilizing the accompanying numerical model wherein the once-over of moths is imparted using the going

X1,1 X1,2 X1,n
X2,1 X2,2 X2,n

X = ⎤

. . . .
. . . . .
. . . . . ⎦

with Eq. 1.

Here m alludes to the no. of moths and n alludes to the aspects. Eq. 2 mirrors the cluster fpr putting away the wellness worth of the moths.

FX 1

FX = (2)

The rundown of flames is displayed in Eq. 3 wherein in a state of harmony with the past conditions x signify the quantity of moths and n shows the amount of angles. Eq. 4 mirrors the group for taking care of the flame wellness values.

Y1,1 Y1,2 Y1,n
Y2,1 Y2,2 Y2,n
. . . .

Y = Ym,1 Ym,2 Ym,n FY 1 (3)

FY = (4)

Here m indicates the quantity of flames. The MFO gives a

three tuple arrangement which evaluates and presents the first arrangement according to the accompanying Eq. 5. XYF = (P, Q, R) (5)

P is a mathematical capacity here that produces sporadic moth test size and well being regard as presented in Eq. 6 and Eq. 7. Q is a fundamental ability which keeps the moths getting across the circle considered. Accepting the action gets fulfilled in Eq. 8, R makes the value 1 or most likely 0 incase the standard isn't met. The upper and lower endpoints of the limits are imparted as a show as portrayed in Eq. 9 and Eq.10.

$P : \emptyset \rightarrow X, FX$ (6)

$Q : X \rightarrow X$ (7)

$R : X \rightarrow 1, 0$ (8)

$Z = [Z1, Z2, Z3, \ldots, Zm-1, Zm]$ (9)

$Z = [z1, z2, z3, \ldots, zm-1, zm]$ (10)

The position of the moth with reference to the flame is presented in Eq. 11. Xi = (Xi, Yj) (11)

Here Xi presents the i th moth, Yi presents the j th fire and Z address winding capabilities. The moth update process is enrolled using a logarithmic winding model. The moths position towards the heading of the flares follows a winding model following the under referred to conditions:

1) The primary characteristic of the twisting should begin with the moth.
2) The end point of the twisting should end with the fire is used to select the most appropriate features.
3) The scope of the winding shouldn't go past or away from the pursuit space which is given in Fig. 3. Z Xi, Yj = Li.gfh. cos (2πh) + Yj (12)

Here Xi mirrors the distance between i th moth and the j th fire. Likewise, f is a value that reflects the condition of the winding and h is an unpredictable number going from [−1, 1] as shown in Eq. 12. The distance between the i th moth and j th is resolved using Eq. 13 as shown underneath: Yi = Yj − Xi (13)

C. VIT MODEL

Spurred by the result of Transformers [33] in machine trans- lation, Without convolution models simply rely upon trans- former layers have coursed around the web in PC vision. In particular, Vision Transformer (ViT) [34] is the really such outline of a transformer-based procedure to match or try and outflank CNNs for image game plan. Various varieties of vision transformers have in like manner been actually suggested that includes refining for data capable planning of vision transformer [35], pyramid structures like CNNs [36], then again self-mindfulness with respect to additionally fos- ter the adequacy through learning a hypothetical depiction rather than playing out all-to-all self consideration [37]. Per- ceiver [38] utilize an astray thought instrument to iteratively distil inputs into a tight lethargic bottleneck, allowing it to scale to manage extraordinarily immense sources of info. Lit- tle portion of ViT based images [24] familiarizes a layer- wise change to encode the huge close by structure for each token as opposed to the guileless tokenization used in ViT [34].

C. PROPOSED EVIT MODEL

Vision Transformer (ViT) [34] initial devotees an image into a progression of fix tokens by parceling it with a specific fix size and a short time later straightforwardly expanding each fix into little fragment. An Additional game plan section (EAS) is added to the course of action, as in the primary MFO result [39]. Furthermore, since self-thought in the Straight projector for smooth images fragments is position agnostic what's more, vision applications extraordinarily

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
www.ijprems.com
editor@ijprems.com
Vol. 04, Issue 04, April 2024, pp: 2582-2590
e-ISSN : 2583-1062
Impact Factor: 5.725

need posi- tion information, ViT adds position embedding into each portion, including the EAS token. A brief time frame later, all tokens are gone through stacked transformer encoders finally the EAS token is used for plan. A transformer encoder is made from a gathering of blocks where each block contains mul- tiheaded self-thought (MSC) with a feed-forward network (FFN). FFN contains two layer multi-layer perceptron with developing extent ’r’ at the mystery layer, and one GELU non-linearity is applied after the essential direct layer. Layer normalization (LS) is applied before each block, and staying substitute ways after each block. The commitment of ViT, x0, moreover, the treatment of the k-th block can be conveyed as,

x0and = [xEAT ||xpatch] + xpos (14) ykand = [xk − 1 + MSC(LN (xk − 1))] (15) xkand = yk + FFN (LN (yk)) (16)

where xcls R1 C and xpatch RN C are the EAS and fix token portions independently and xpos R(1 N) C is the position introducing. N and C are the amount of fix token sections what’s more, part of the introducing, independently. It is crucial to note that the EAS token is one extremely unexpected ViT design from CNNs. In CNNs, the last encoding is primarily obtained by combining the components across all spatial areas, whereas in ViT, the last encoding is the EAS associated with the fix tokens at each transformer encoder. In like manner, we consider level images as an expert that summarizes all the fix tokens and consequently the proposed module is arranged considering smooth bug image to shape the proposed Enhanced Vision transformer architecture (EViTA). The total work process of proposed model is shown in Fig. 4.

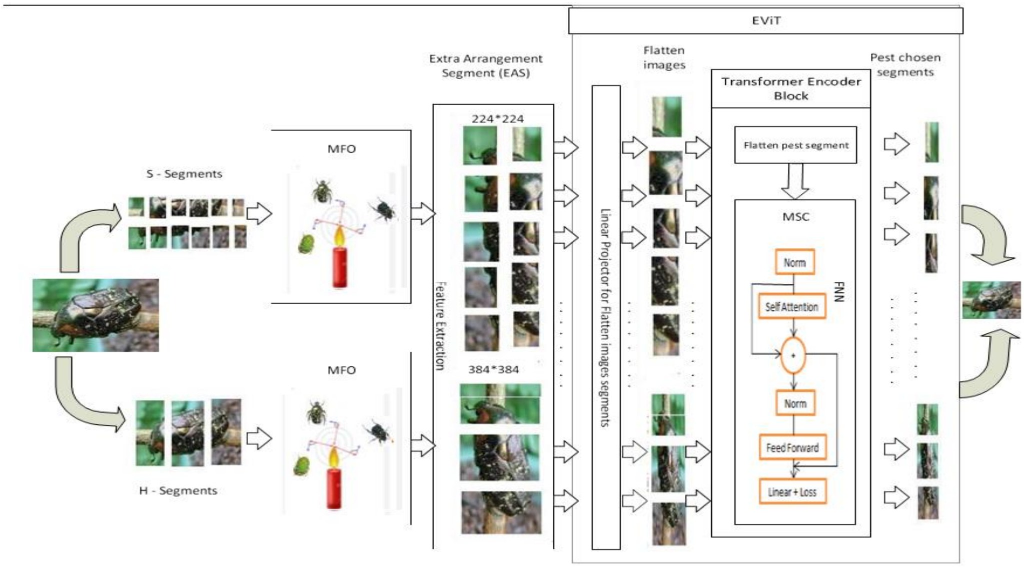

FIGURE 4. Proposed Enhanced Vision Transformer Architecture (Evita) Model.

3. RESULTS AND DISCUSSION

The collected datasets are carried out in to Google colab environment for experiment, and for data analysis purpose Python 3.7 programming language is used. The aggregationof the fix size influences the precision and diverse plan of ViT with fine-grained fix size, ViT can perform far supe- rior accomplishes higher Disillusionments and memory use.A ViT with a fix size of 16 overcomes a ViT with a fix size of 32 by 6%, but the former requires 4 more tumbles. This has prompted us to suggest an approach that aims to changethe complexity while utilizing the advantages of much more finely tuned fix sizes. Much more unambiguously, we ini- tially presented a twofold separating ViT where each branchoperates at a different scale (or fix size in the fix implanting),and then we quickly follow that up by proposing a substan- tial yet persuasive module to consolidate data between the branches. Fig. 4 frames the affiliation plan of our proposed Enhanced vision transformer learning design (EViTA). Our model is generally produced using K multiscale transformerencoders where each encoder incorporates two branches:

(1) H-Fragment: a monstrous (principal) branch that uses coarsegrained fix size (Pl) with extra transformer encoders and more prominent installing perspectives,

(2) S-Portion:to some degree (complementary) branch that works at fine- grained fix size (Ps) with not such a lot of encoders yet rathermore unpretentious inserting viewpoints. The two branches are consolidated L times and the EAS recognizable proofof the two branches around the end are utilized for suppo- sition. Be aware that we additionally add a learnable posi- tion implanting before the multiscale transformer encoder forlearning position data as in ViT [34].

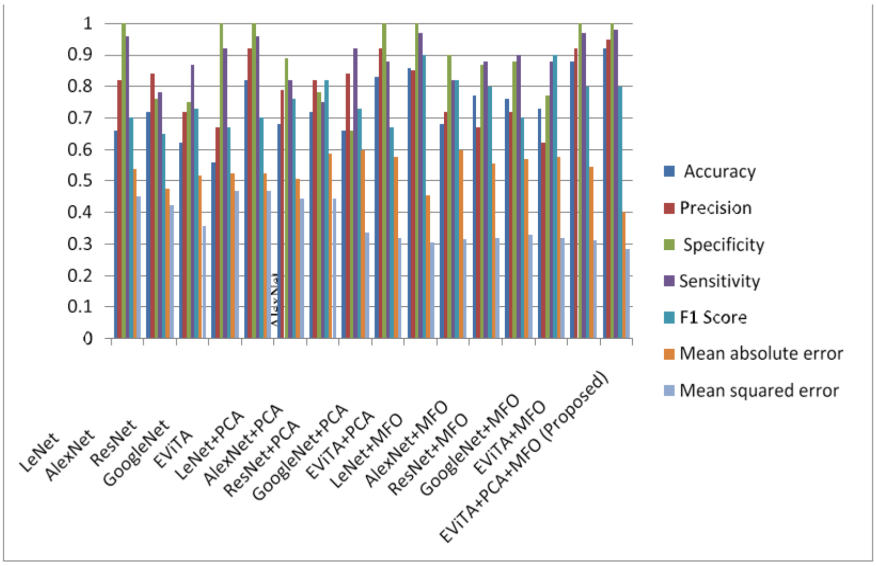

FIGURE 5. Proposed performance analysis chart of proposed EViTA model.

TABLE 2. Overall pest images versus predicted pest images.

Overall pest images	**Predicted pest images**													
				GoogleNet		LeNet with PCA	AlexNet with PCA	ResNet with PCA	GoogleNet with PCA	LeNet with MFO	AlexNet with MFO	ResNet with MFO	GoogleNet with MFO	EViTA with PCA + MFO
Aphids Training DS	6	6	6	67	5	6	6	6	66	5	5	6	62	6
Aphids Testing DS	6	6	6	62	5	6	5	5	62	5	5	5	54	5
Wireworm Training DS	6	6	7	75	7	6	6	7	68	8	7	7	75	7
Wireworm Testing DS	6	6	6	69	6	6	6	7	74	6	6	6	67	6
Gram Caterpillar Training I	6	6	6	68	6	6	6	6	66	5	6	5	64	6
Gram Caterpillar Testing D	6	6	6	65	7	6	6	6	67	7	6	6	70	7

The essential ViT [34] accomplishes serious outcomes showed up distinctively according to likely the best CNN models yet right when prepared for exceptionally immense degree datasets (for example Aphids [28] and Gramcater- pillar [29]). It just so happens, Gram caterpillar [29] shows that with the assistance of a rich arrangement of information improvement strategies, ViT can be prepared from ImageNetin MFO alone to pass practically identical outcomes on to CNN models. Fittingly, in our assessments, we gather our models considering [39], and apply their default hyperlimits for arranging. During assessment, we resize the more con- fined side of an image to 256 and take the middle harvest 224 224 as the information. Besides, we likewise adjustedour models to a more prominent goal (384 384) for fair relationship on occasion. Bicubic consideration was appliedto change the size of the learnt position inserting, and the finetuning took 30 insect images. More subtleties ought to be apparent as in bracing material [35].

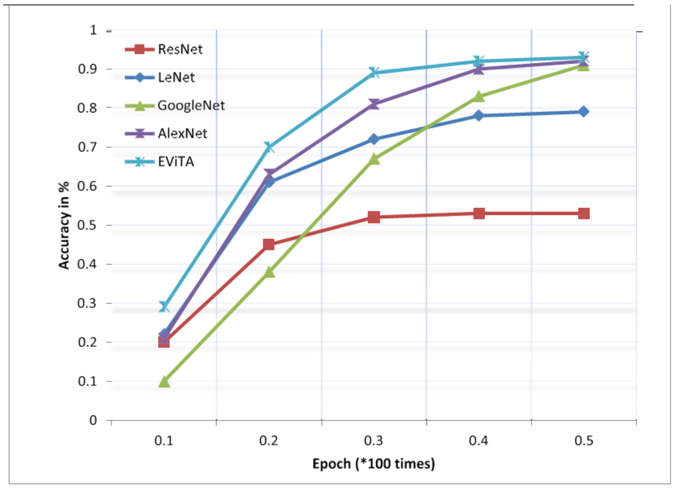

FIGURE 6. Accuracy rate with different epochs.

The gathered Aphids (IP102 Dataset), Wireworm (IP102 Dataset), and Gram Caterpillar was taken from Kaggle data set are executed in Google colab. The datasets are having higher inconstancy required preprocessing. Toward the beginning, the missing characteristics were managed using the property mean technique wherein the they were finished offwith the mean of attribute values. It's certainly a fact that CNN can be executed solely on numerical dataset values.

TABLE 3. Performance analysis of proposed model versus existing models

Model used	Accur	Precis	Specifi	Sensiti	F1 Sc	Mean absolute er	Mean squared er
LeNet	.66	.82	1.0	.96	0.7	5.38	45.23
	.72	.84	.76	.78	0.65	4.76	42.2
	.62	.72	.75	.87	0.73	5.17	35.70
	.56	.67	1.0	.92	0.67	5.25	46.69
	.82	.92	1.0	.96	0.7	5.25	46.69
LeNet+PCA	.68	.79	.89	.82	0.76	5.08	44.49
AlexNet+PCA	.72	.82	.78	.75	0.82	5.88	44.45
ResNet+PCA	.66	.84	.66	.92	0.73	5.98	33.56
GoogleNet+PCA	.83	.92	1.0	.88	0.67	5.76	31.78
	.86	.85	1.0	.97	0.9	4.55	30.34
	.68	.72	.9	.82	0.82	5.98	31.33
AlexNet+MFO	.77	.67	.87	.88	0.8	5.55	31.78
ResNet+MFO	.76	.72	.88	.9	0.7	5.68	32.86
GoogleNet+MFO	.73	.62	.77	.88	0.9	5.78	31.78
EViTA+MFO	.88	.92	1.0	.97	0.8	5.45	31.25
EViTA+PCA+MFO (Propos	.92	.95	1.0	.98	0.8	4.01	28.36

The present dataset having non-numerical characteristicsshould be presented to One Hot Encoding technique to ensureits likeness to be dealt with into the CNN model. Thisstandardized and pre-handled data was moreover normalized using StandardScaler strategy. StandardScaler normalizationtechnique added to fit in all potential gains of the charac- teristics to a particular scale. Then, at that point, inorder topick the ideal components that would basically influence the class marks (real insect aftection), MFO estimation was used.This decreased dataset was then dealt with into a CNN model for expectation of influenced nut leafs. An amount of 70%of the dataset was used to set up the CNN models and the rest was used for testing the model. The after effects of theproposed MFO PCA EViTA model were at last thought

about in contrast to the cutting edge conventional CNNmodels in particular LeNet, AlexNet, ResNet, GoogleNet, EViTA, LeNet PCA, AlexNetPCA, ResNet PCA,GoogleNet PCA, EViTAPCA, LeNetMFO, AlexNetMFO, ResNetMFO, GoogleNetMFO, EViTA MFO thinking about the measurements - exactness, accuracy,awareness, explicitness, F1 score, mean outright mistake, andmean squared blunder in Fig. 5. With a goal to pick the idealboundaries like number of layers, improvement capability,actuation capability, and so on for the CNN model, latticesearch

Fig. 6 portrays the expectation results achieved by LeNet,AlexNet, ResNet, GoogleNet, EViTA, LeNet with PCA,

AlexNetPCA, ResNetPCA, GoogleNet PCA,EViTAPCA, LeNetMFO, AlexNetMFO, ResNet MFO, GoogleNetMFO, EViTAMFO and its variations.It is apparent from the figure that EViTA with MFO PCAyielded enhanced expectation brings about correlation withother different variations of CNN.

Likewise Fig.+6 portrays the outcomes acqu+ired with var-ious ages of LeNet, GoogleNet, ResNet, AlexNet, and our proposed EViT+A. It is further apparent+from this figure that

+CNN calculations, in mix with MFO for highlight extractionhave beated different partners. It could hence be laid out thatthe proposed EViTA PCA MFO has been fruitful in creat- ing improved brings about correlation with different models

thought about in the review Table. 2.

Sums up the outcomes achieved by every one of the thought about models. Table. 3 features the enhanced exhibition of the proposed model against the other best in class mod-els for the measurements exactness, accuracy, explicitness, responsiveness, F1 score, mean outright mistake, and mean squared blunder. The outcomes additionally finish up the way that considering the evaluation measurements values inassessment of the worldwide optima, EViTA strategy has been fruitful in accomplishing better forecast results using essentially less preparation time.

4. CONCLUSION

This work presented an EViTA, the two layer vision trans- former pest image segments for learning multi-scale high-lighted parameters, to further develop the acknowledgment accuracy for image order. Using MFO the fetures of chosen pest images are extracted and dumbed into the EAS block,it further foster a combination technique in light of cross- regard for trade data between two branches proficiently in direct time. With broad investigations, that's what we exhibitour proposed model performs better compared to or compa- rable to a few simultaneous deals with vision transformer, notwithstanding effective CNN models. The present review has utilized the Aphids (IP102 Dataset), Wireworm (IP102 Dataset), and Gram Caterpillar datasets gathered from the public accessible vault. Choice of the ideal elements from this dataset goes about as an essential errand of the review sincethe flatten the images using linear projection method. The unmitigated qualities of the chosen dataset is switched over completely to mathematical qualities utilizing the One hot encoding method and the StandardScaler strategy is carried out for information standardization. The ideal highlights in the dataset are chosen utilizing the MFO calculation, which are taken care of into the CNN model for pest image pre- diction. While our proposed EVITA model starts to expose what's underneath on dual branch vision transformers for image portrayal, we surmise that in future there will be further works in making powerful multi branch transformers for other vision applications, including object revelation, seman- tic division, and video movement affirmation.

5. REFERENCES

[1] H. Qi, B. Zhu, L. Kong, W. Yang, J. Zou, Y. Lan, and L. Zhang, ''Hyper-spectral inversion model of chlorophyll content in peanut leaves,'' Appl. Sci., vol. 10, no. 7, p. 2259, Mar. 2020.

[2] H. Qi, Y. Liang, Q. Ding, and J. Zou, ''Automatic identification of peanut- leaf diseases based on stack ensemble,'' Appl. Sci., vol. 11, no. 4, p. 1950, Feb. 2021.

[3] C. R. Rahman, P. S. Arko, M. E. Ali, M. A. I. Khan, S. H. Apon, F. Nowrin, and A. Wasif, ''Identification and recognition of Rice diseases and pests using convolutional neural networks,'' Biosyst. Eng., vol. 194, pp. 112–120, Jun. 2020.

[4] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, ''Identification of Rice diseasesusing deep convolutional neural networks,'' Neurocomputing, vol. 267, pp. 378–384, Dec. 2017.

[5] B. B. Sharma, G. Gupta, P. Vaidya, S. Basheer, F. H. Memon, and R. N. Thakur, ''Internet of Things-based crop classification model using deep learning for indirect solar drying,' Wireless Commun. Mobile Com- put., vol. 2022, pp. 1–11, Jun. 2022.

[6] R. M. Rakholia, J. H. Tailor, J. R. Saini, J. Kaur, and H. Pahuja, ‘‘Ground- nuts leaf disease recognition using neural network with progressive resiz- ing,’’ Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 6, pp. 1–6, 2022.

[7] M. Konate, J. Sanou, A. Miningou, D. K. Okello, H. Desmae, P. Janila, and R. H. Mumm, ‘‘Past, present and future perspectives on groundnut breedingin Burkina Faso,’’ Agronomy, vol. 10, no. 5, p. 704, May 2020.

[8] A. Tomar, G. Gupta, W. Salehi, C. Vanipriya, N. Kumar, and B. Sharma, ‘‘A review on leaf-based plant disease detection systems using machine learning,’’ in Proc. Recent Innov. Comput. (ICRIC), vol. 1. Jammu and Kashmir, India: Springer, May 2022, pp. 297–303.

[9] K. Thenmozhi and U. S. Reddy, ‘‘Crop pest classification based on deep convolutional neural network and transfer learning,’’ Comput. Electron. Agricult., vol. 164, Sep. 2019, Art. no. 104906.

[10] M. E. Karar, F. Alsunaydi, S. Albusaymi, and S. Alotaibi, ‘‘A new mobile application of agricultural pests recognition using deep learning in cloud computing system,’’ Alexandria Eng. J., vol. 60, no. 5, pp. 4423–4432, Oct. 2021.

[11] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, ‘‘Recent advances in image processing techniques for automated leaf pest and disease recognition—A review,’’ Inf. Process. Agricult., vol. 8, no. 1, pp. 27–51, Mar. 2021.

[12] M. Tariq, S. Nisar, A. Shah, S. Akbar, M. A. Khan, and S. Z. Khan, ‘‘Effect of hybrid reinforcement on the performance of filament wound hollow shaft,’’ Compos. Struct., vol. 184, pp. 378–387, Jan. 2018.

[13] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, ‘‘SLICsuperpixels compared to state-of-the- art superpixel methods,’’ IEEE Trans.Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[14] T. F. Chan and L. A. Vese, ‘‘Active contours without edges,’’ IEEE Trans. Image Process., vol. 10, no. 2, pp. 266–277, Feb. 2001.

[15] M. Islam, A. Dinh, K. Wahid, and P. Bhowmik, ‘‘Detection of potato dis- eases using image segmentation and multiclass support vector machine,’’ in Proc. IEEE 30th Can. Conf. Electr. Comput. Eng. (CCECE), Apr. 2017,pp. 1–4.

[16] D. P. Hughes and M. Salathe, ‘‘An open access repository of images on plant health to enable the development of mobile disease diagnostics,’’ 2015, arXiv:1511.08060.

[17] P. Zhang, X. Dai, J. Yang, B. Xiao, L. Yuan, L. Zhang, and J. Gao, ‘‘Multi-scale vision longformer: A new vision transformer for high-resolution image encoding,’’ in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2021, pp. 2978–2988.

[18] L. Yuan, Y. Chen, T. Wang, W. Yu, Y. Shi, Z. Jiang, F. E. H. Tay, J. Feng, and S. Yan, ‘‘Tokens-to-token ViT: Training vision transformers from scratch on ImageNet,’’ in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV),Oct. 2021, pp. 538–547.

[19] F. Almalik, M. Yaqub, and K. Nandakumar, ‘‘Selfensembling vision trans- former (SEViT) for robust medical image classification,’’ in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Singapore: Springer,Sep. 2022, pp. 376–386.

[20] A. Wang, S. Xing, Y. Zhao, H. Wu, and Y. Iwahori, ‘‘A hyperspectral image classification method based on adaptive spectral spatial kernel combined with improved vision transformer,’’ Remote Sens., vol. 14, no. 15, p. 3705,Aug. 2022.

[21] C. R. Chen, Q. Fan, and R. Panda, ‘‘CrossViT: Cross-attention multi-scale vision transformer for image classification,’’ in Proc. IEEE/CVF Int. Conf.Comput. Vis. (ICCV), Oct. 2021, pp. 347–356.