# **Environments**<br> *Abstract*—Infacility-based agricultural environments,<br> *Abstract*—In facility-based agricultural environments,<br> *Abstract*—In facility-based agricultural environments,<br> *Abstract*—In facility-based agricu **accurately identifying green tomatoes presents a significant**<br> **accurately identifying green tomatoes presents a significant**<br> *Abstract*—In facility-based agricultural environments, areas and production<br>
accurately ident tation Model Based on<br>Igorithm Under Facility<br>imgyu Yan, Weikuan Jia<br>fingyu Yan, Weikuan Jia<br>nutritional content. With the continuous expansion of planted<br>areas and production, the case of transportation and storage<br>of gre tation Model Based on<br>Igorithm Under Facility<br>Singyu Yan, Weikuan Jia<br>nutritional content. With the continuous expansion of planted<br>areas and production, the case of transportation and storage<br>of green tomatoes further enh Engineering Letters<br>Green Tomato Segmentation Model Based on<br>Optimized Swin-Unet Algorithm Under Facility<br>Environments<br>Ru Jiang, Huichuan Duan, Jingyu Yan, Weikuan Jia Engineering Letters<br>
Green Tomato Segmentation Model Based on<br>
Optimized Swin-Unet Algorithm Under Facility<br>
Environments<br>
Ru Jiang, Huichuan Duan, Jingyu Yan, Weikuan Jia<br>
The Lagorithm Lagorithm (Still Me Continuous expa Environments Engineering Letters<br>
ato Segmentation Model Based on<br>
Swin-Unet Algorithm Under Facility<br>
Environments<br>
Ru Jiang, Huichuan Duan, Jingyu Yan, Weikuan Jia<br>
mutritional content. With the continuous expansion of planted<br>
gricu

**challenge for machine vision systems due to the color similarly due to the color similarly dentifying green tomatoes presents a significant particional content. We accurately identifying green tomatoes presents a signific between from the fruits and background branches and leaves as and production.**<br> **between facility-based** agricultural environments, areas and production, the ease<br>
accurately identifying green tomatoes presents a signific **ENVITONIMENTS**<br> **ENVITONIMENTS**<br> **WELD ABOUT ASSEM**<br> **WELD ABOUT ASSEM**<br> **WELD ABOUT AND THE MULTIME CONTINUMERT AND MENTAIN THE SECAULT AND THE SECAULT AND THE SECAULT AND HE SECAULT AND WELL AS the overlapping occlusion Problem, the study constructs and political constrants and problem from the frequency of the problem, the frequency of the Attention States and production, the accurately identifying green tomatoes presents a significant (AG)** module using Swin-Unet as the baseline model, or the difference of the difference the original environments, areas and production, the execurately identifying green tomatoes presents a significant of green tomatoes Ru Jiang, Huichuan Duan, Jingyu Yan, Weikua:<br> *Mbstract*—In facility-based agricultural environments,<br>
areas and production,<br>
accurately identifying green tomatoes presents a significant<br>
of green tomatoes furt<br>
challenge **suppress irrelevant regions in the background, and effectively ENTRACE THE SET CONSIGNATISE SET ALL ASSEM CONDUCT AND EXERCT AND A SERVIDE SERVIDENT (AG) module using Swin-Unet as the baseline model, so t** *Abstract***—In facility-based agricultural environments, areas and production, the accurately identifying green tomatoes presents a significant of green tomatoes furthe between green fruits and background branches and leav Example 1 set to the study in the study in the study in the study in the content of green tomatoes further between green fruits and background branches and leaves as inficiant between green fruits and background branche** *Abstract***—In facility-based agricultural environments,** areas and production, execurately identifying green tomatoes presents a significant of green tomatoes further challenge for machine vision systems due to the color **Expanding the set allow the set allenge for mather vision systems due to the color similarity and between green fruits and background branches and leaves as challeng** acturately userlary intervigy green tomatoes presents a signment<br>the challenge for machine vision systems due to the color similarity<br>between green fruits and background branches and leaves as<br>well as the overlapping occlu Experimental results on the specially constrained by the equilibration<br>
between green fruits and background branches and leaves as<br>
well as the overlapping occlusion between fruits. To solve this<br>
problem, this study const **Drevively** and **Drevively** and **Drevively** and **exactly out and the model show that the model as the overlapping occlusion between fruits. To solve this face numerous challenges in the controller, this study constructs an** We as the overapping occusion between intents. To sove the search and palmanent and palmanent and the phase and the search actes of the search are the presentation of target features. Additionally, in enhance the represent **INCRET TERM**<br> **INCRED TO THE TRANSM**<br> **INCRED TO THE TRANSM**<br> **INCRED TO THE TRANSM**<br> **INCRED TO THE TRANSMIS INTERENT THE TRANSMISE SUPPLEM IN THE TRANSMISE SUPPLEM IN THE TRANSMISE SUPPRESS irrelevant regions in the bac** (AST) module using symm-there as the baskene mondels, but all the matter in the same the representation of target features. Additionally, in subpress irrelevant regions in the background, and effectively picking robots a f moute can locus on the reatures renated to green tomatoes, have a foculate the representation of target features. Additionally, in automatic picking invorder to optimize the edge smoothing of green tomato a chrous Spatial suppless irrecevant regions in the background, and encertained responses the representation of target features. Additionally, in order to optimize the edge smoothing of green tomato segmentation, this study further introdu entate the representation of range teatures. Additional and entergies and entergies and the segmentation, this study further introduces a Atrous Spatial using computer vision [4-Pyramid Pooling (ASPP) module, which signifi or<br>
or the riginal subseque subcounting of green formation, this study further introduces a Atrous Spatial<br>
Pyramid Pooling (ASPP) module, which significantly improves struggle to recognize green fr<br>
the segmentation accur **Expirimation, this starty in the filtrowner into uncered as Atolus Spatial Pooling (ASPP) module, which significantly improves** struggle to recognize green frithe segmentation accuracy by expanding the feature extraction **Experimentation aroung (ASPT) module, when significantly improves<br>
the segmentation accuracy by expanding the feature sensing Similarity with the backgroof<br>
deld and enhancing the multi-scale feature extraction capability** Field and enhancing the multi-scale feature extraction easing<br> **Field and enhancing the multi-scale feature extraction capability.** Occlusion among freen Experimental results on the specially constructed green tomato<br>
for **incularing the muni-scare leading to the sparing in the sparing the muni-scare extraction capability contributed green to semantic segmentation** dataset show that the model achieves 97.5%, 92.4% and 85.9% for Pixel Accura **prospect. IDENTE ACTE ACT ACTS AND EXTENDINICTION**<br> **ICHOT ACT TEXEL ACT THE ACT AND THE SET AND THE STAR AND THE** Final Accuracy (FA), Dice similarity coencelent (Dice)<br>Intersection over Union (IOU), respectively. The new moutperforms existing partial semantic segmentation model<br>several key metrics, proving its effectiveness in comple partial semantic segmentation models in<br>
roving its effectiveness in complex facility<br>
search not only addresses the technical<br>
targ green fruits, but also provides solid<br>
r the development and application of<br>
lequipment. difficulties in recognizing green fruits, but also provides solid<br>
technical support for the development and application of<br>
intelligent agricultural equipment. The model can be applied to<br>
intenting fertilizer waste. In<br>

<span id="page-0-0"></span>

From Summaton, ASPP, AG<br>
I. INTRODUCTION<br>
I. INTRODU

(2021GXRC049).<br>R.Jiang is a postgraduate student of School of Information Science and Engineering, Shandong Normal University, Jinan (superjr1102@163.com);

hcduan@sdnu.edu.cn);<br>J.Y.Yan is a postgraduate student of School of Information Science and (1543670705@qq.com)<br>W.K. Jia is an associate professor of School of Information Science and

(180.: 2022KJ250); New Twentueth Trems of Universities in Jin<br>
(2021GXRCO49). Response a postgraduate student of School of Information Science a<br>
Engineering, Shandong Normal University, Jinan 250358, Chi<br>
(superjr1102@163 (2021GXRCO49).<br>
R.Jiang is a postgraduate student of Scl<br>
Engineering, Shandong Normal Unive<br>
(superjr1102@163.com);<br>
H.C. Duan is a professor of School of Info<br>
Shandong Normal University, Jinan 25035<br>
heduan@sdnu.edu.cn)

**Such the Facility**<br> **Consider Facility**<br> **Consider Facility**<br> **Consider Facility**<br> **Consider Facility**<br> **Consider a** must<br> **Consider a** market appeal [1-2].<br> **However, the cultivation and management of green tomatoes**<br> **f** 1gorithm Under Facility<br>
ingyu Yan, Weikuan Jia<br>
nutritional content. With the continuous expansion of planted<br>
areas and production, the ease of transportation and storage<br>
of green tomatoes further enhances their market **if**<br> **ingyu Yan, Weikuan Jia**<br> **Singyu Yan, Weikuan Jia**<br> **nutritional content.** With the continuous expansion of planted<br>
areas and production, the ease of transportation and storage<br>
of green tomatoes further enhances their m **ITTICTILS**<br>
ingyu Yan, Weikuan Jia<br>
nutritional content. With the continuous expansion of planted<br>
areas and production, the ease of transportation and storage<br>
of green tomatoes further enhances their market appeal [1-2] lingyu Yan, Weikuan Jia<br>
nutritional content. With the continuous expansion of planted<br>
areas and production, the ease of transportation and storage<br>
of green tomatoes further enhances their market appeal [1-2].<br>
However, Fingyu Yan, Weikuan Jia<br>
nutritional content. With the continuous expansion of planted<br>
areas and production, the ease of transportation and storage<br>
of green tomatoes further enhances their market appeal [1-2].<br>
However, Fingyu Yan, Weikuan Jia<br>
nutritional content. With the continuous expansion of planted<br>
areas and production, the ease of transportation and storage<br>
of green tomatoes further enhances their market appeal [1-2].<br>
However, nutritional content. With the continuous expansion of planted<br>areas and production, the ease of transportation and storage<br>of green tomatoes further enhances their market appeal [1-2].<br>However, the cultivation and manageme nutritional content. With the continuous expansion of planted<br>areas and production, the ease of transportation and storage<br>of green tomatoes further enhances their market appeal [1-2].<br>However, the cultivation and manageme nutritional content. With the continuous expansion of planted<br>areas and production, the ease of transportation and storage<br>of green tomatoes further enhances their market appeal [1-2].<br>However, the cultivation and manageme areas and production, the ease of transportation and storage<br>of green tomatoes further enhances their market appeal [1-2].<br>However, the cultivation and management of green tomatoes<br>face numerous challenges in the complex e of green tomatoes further enhances their market appeal [1-2].<br>However, the cultivation and management of green tomatoes<br>face numerous challenges in the complex environment of<br>facility-based agriculture [3]. Among these cha However, the cultivation and management of green tomatoes<br>face numerous challenges in the complex environment of<br>facility-based agriculture [3]. Among these challenges,<br>harvesting is particularly time-consuming and<br>labor-i face numerous challenges in the complex environment of facility-based agriculture [3]. Among these challenges, harvesting is particularly time-consuming and labor-intensive, making the development of automatic picking robo facility-based agriculture [3]. Among these challenges,<br>harvesting is particularly time-consuming and<br>labor-intensive, making the development of automatic<br>picking robots a focal point of interest. The primary task of<br>autom harvesting is particularly time-consuming and<br>labor-intensive, making the development of automatic<br>picking robots a focal point of interest. The primary task of<br>automatic picking involves accurately locating the fruits<br>usi labor-intensive, making the development of automatic<br>picking robots a focal point of interest. The primary task of<br>automatic picking involves accurately locating the fruits<br>using computer vision [4-6]. Traditional machine picking robots a focal point of interest. The primary task of automatic picking involves accurately locating the fruits using computer vision [4-6]. Traditional machine methods struggle to recognize green fruits accurately automatic picking involves accurately locating the fruits<br>using computer vision [4-6]. Traditional machine methods<br>struggle to recognize green fruits accurately due to their color<br>similarity with the background foliage and using computer vision [4-6]. Traditional machine methods struggle to recognize green fruits accurately due to their color similarity with the background foliage and the overlapping occlusion among fruits [7]. Therefore, th struggle to recognize green fruits accurately due to their color<br>similarity with the background foliage and the overlapping<br>occlusion among fruits [7]. Therefore, the use of advanced<br>semantic segmentation techniques to acc similarity with the background foliage and the overlapping<br>occlusion among fruits [7]. Therefore, the use of advanced<br>semantic segmentation techniques to accurately segment and<br>localize green tomatoes is of paramount impor occlusion among fruits [7]. Therefore, the use of advanced<br>semantic segmentation techniques to accurately segment and<br>localize green tomatoes is of paramount importance. This<br>technique not only contributes to efficient agr semantic segmentation techniques to accurately segment and localize green tomatoes is of paramount importance. This technique not only contributes to efficient agricultural management and planning but also optimizes resour localize green tomatoes is of paramount importance. This<br>technique not only contributes to efficient agricultural<br>management and planning but also optimizes resource<br>allocation. For instance, water sources can be more prec technique not only contributes to efficient agricultural<br>management and planning but also optimizes resource<br>allocation. For instance, water sources can be more precisely<br>targeted to tomato-growing areas, significantly red harvesters. ocation. For instance, water sources can be more precisely<br>geted to tomato-growing areas, significantly reducing<br>ter wastage. Similarly, fertilizer application can be<br>justed according to the distribution of tomatoes, effec targeted to tomato-growing areas, significantly reducing<br>water wastage. Similarly, fertilizer application can be<br>adjusted according to the distribution of tomatoes, effectively<br>minimizing fertilizer waste. In automated har water wastage. Similarly, fertilizer application can be adjusted according to the distribution of tomatoes, effectively minimizing fertilizer waste. In automated harvesting, robots equipped with semantic segmentation capab adjusted according to the distribution of tomatoes, effectively<br>minimizing fertilizer waste. In automated harvesting, robots<br>equipped with semantic segmentation capabilities can detect<br>and locate fruits, thereby reducing r

**EXECUTE SURFACE SET AND SET A** From External University<br>
Tomatoes are highly valued in agriculture due to their rich<br>
In recent years<br>
as a prominent<br>
Tomatoes are highly valued in agriculture due to their rich<br>
This work is supported by National Valued I. INTRODUCTION<br>
I. INTRODUCTION<br>
In recent years, im.<br>
as a prominent foculation of descending the efficiency of the segmentation of descending the segmentation of descending the segmentation technic<br>
Manuscript received I. INTRODUCTION<br>
In recent years, imag<br>
as a prominent focus<br>
strategradion of deep<br>
segmentation techniqu<br>
Manuscript recived April 17, 2024; revised September 24, 2024.<br>
This work is supported by National Nature Science Inrecent years, im<br>
as a prominent focus are highly valued in agriculture due to their rich<br>
Manuscript received April 17, 2024, revised September 24, 2024.<br>
This work is supported by National Nature Science Foundation of omatoes are highly valued in agriculture due to their rich<br>
segmentation of<br>
segmentation of<br>
Manuscript received April 17, 2024; revised September 24, 2024.<br>
This work is supported by National Nature Science Foundation of I omatoes are highly valued in agriculture due to their rich<br>
segmentation techn<br>
Manuscript received April 17, 2024, revised September 24, 2024.<br>
This work is supported by National Nature Science Foundation of China (No.: Manuscript received April 17, 2024; revised September 24, 2024.<br>
This work is supported by National Nature Science Foundation of China (No.:<br>
27289); Young Imovation Them Program of Shandong Provincial University<br>
0.: 2022 Manuscript received April 17, 2024, revised September 24, 2024.<br>
This work is supported by National Nature Science Foundation of China (No.:<br>
20072289); Young Innovation Team Program of Shandong Provincial University<br>
(202 Ins work is supported by Natuonal Nature Science roundation of China (No.: 2022KJ250); New Twentieth Items of Universities in Jinan particularly when the D21GXRC049).<br>
20122KJ250); New Twentieth Items of Universities in Ji overlay to dig movador ream rogram of Shandong Provincial University<br>
(No.: 2022K1250); New Twentieth Items of Universities in Jinan particularly v<br>
(2021GXRC049).<br>
R.Jiang is a postgraduate student of School of Informatio minimizing fertilizer waste. In automated harvesting, robots<br>equipped with semantic segmentation capabilities can detect<br>and locate fruits, thereby reducing reliance on human labor<br>[8]. Thus, enhancing the visual system's equipped with semantic segmentation capabilities can detect<br>and locate fruits, thereby reducing reliance on human labor<br>[8]. Thus, enhancing the visual system's segmentation<br>precision is paramount not only for the effectiv and locate fruits, thereby reducing reliance on human labor [8]. Thus, enhancing the visual system's segmentation precision is paramount not only for the effective management of green tomato cultivation but also as a pivot [8]. Thus, enhancing the visual system's segmentation precision is paramount not only for the effective management of green tomato cultivation but also as a pivotal element in augmenting the efficacy of robotic fruit and v precision is paramount not only for the effective management<br>of green tomato cultivation but also as a pivotal element in<br>augmenting the efficacy of robotic fruit and vegetable<br>harvesters.<br>In recent years, image semantic s of green tomato cultivation but also as a pivotal element in<br>augmenting the efficacy of robotic fruit and vegetable<br>harvesters.<br>In recent years, image semantic segmentation has emerged<br>as a prominent focus in the field of augmenting the efficacy of robotic fruit and vegetable<br>harvesters.<br>In recent years, image semantic segmentation has emerged<br>as a prominent focus in the field of deep learning. The<br>integration of deep learning-based image s harvesters.<br>
In recent years, image semantic segmentation has emerged<br>
as a prominent focus in the field of deep learning. The<br>
integration of deep learning-based image semantic<br>
segmentation techniques with agricultural a In recent years, image semantic segmentation has emerged<br>as a prominent focus in the field of deep learning. The<br>integration of deep learning-based image semantic<br>segmentation techniques with agricultural applications has<br> as a prominent focus in the field of deep learning. The<br>integration of deep learning-based image semantic<br>segmentation techniques with agricultural applications has<br>gradually evolved into a significant area of development integration of deep learning-based image semantic<br>segmentation techniques with agricultural applications has<br>gradually evolved into a significant area of development [9].<br>Studies show that image semantic segmentation offer segmentation techniques with agricultural applications has gradually evolved into a significant area of development [9]. Studies show that image semantic segmentation offers a significant advantage in accurately segmenting gradually evolved into a significant area of development [9].<br>Studies show that image semantic segmentation offers a<br>significant advantage in accurately segmenting fruit targets,<br>particularly when there is a pronounced col Studies show that image semantic segmentation offers a significant advantage in accurately segmenting fruit targets, particularly when there is a pronounced color difference between the fruit and the background. Häni [10]

**Engineering Letters**<br>detection accuracy under ordinary conditions, it was unable among green leaves<br>to completely recognize all the fruits in images taken in demonstrating the effectiver<br>complex environments. Barth et al. **Engineering Letters**<br>detection accuracy under ordinary conditions, it was unable<br>to completely recognize all the fruits in images taken in<br>demonstrating the effectiveness complex environments. Barth et al. [12] used DeepL **Engineering Letters**<br>detection accuracy under ordinary conditions, it was unable<br>to completely recognize all the fruits in images taken in<br>demonstrating the effectiveness complex environments. Barth et al. [12] used DeepL **Engineering Letters**<br>detection accuracy under ordinary conditions, it was unable<br>to completely recognize all the fruits in images taken in<br>demonstrating the effectivenes<br>complex environments. Barth et al. [12] used DeepLa **Engineering Letters**<br>detection accuracy under ordinary conditions, it was unable<br>architecture architecture for each of example the DaSNet-V2 network architecture for real-time<br>proposed the DaSNet-V2 network architecture f **Engineering Letters**<br>detection accuracy under ordinary conditions, it was unable among green leaves<br>to completely recognize all the fruits in images taken in demonstrating the effective<br>complex environments. Barth et al. **Engineering Letters**<br>
detection accuracy under ordinary conditions, it was unable<br>
to completely recognize all the fruits in images taken in demonstrating the effective<br>
complex environments. Barth et al. [12] used DeepLa **Engineering Letters**<br>
detection accuracy under ordinary conditions, it was unable<br>
to completely recognize all the fruits in images taken in demonstrating the effective<br>
complex environments. Barth et al. [12] used DeepL **Engineering Letters**<br>detection accuracy under ordinary conditions, it was unable among green leaves<br>to completely recognize all the fruits in images taken in demonstrating the effective<br>complex environments. Barth et al. **Experimental results demonstrate that the optimal model series are demonstrated results and the optimal model of the optimal model of** detection accuracy under ordinary conditions, it was unable among green leaves to completely recognize all the fruits in images taken in demonstrating the effectivene complex environments. Barth et al. [12] used DeepLab to detection accuracy under ordinary conditions, it was unable among green leaves<br>to completely recognize all the fruits in images taken in demonstrating the effective<br>complex environments. Barth et al. [12] used DeepLab to m detection accuracy under ordinary conditions, it was unable<br>
among green leav<br>
to completely recognize all the fruits in images taken in<br>
demonstrating the effection<br>
segment bell peoper fruits and plants. Kang et al. [13] to completely recognize all the fruits in images taken in demonstrating the effect<br>complex environments. Barth et al. [12] used DeepLab to more stable and less<br>segment bell pepper fruits and plants. Kang et al. [13] Altho complex environments. Barth et al. [12] used DeepLab to more stable and less<br>segment bell pepper fruits and plants. Kang et al. [13] Although deep learning<br>proposed the DaSNet-V2 network architecture for real-time image se segment bell pepper fruits and plants. Kang et al. [13] Although deep learning has<br>proposed the DaSNet-V2 network architecture for real-time image segmentation, applicatie<br>detection and semantic segmentation of apples and proposed the DaSNet-V2 network architecture for real-time<br>detection and semantic segmentation of apples and branches<br>in orchard environments using visual sensors. This network inherent challenges, such<br>enhances feature ext detection and semantic segmentation of apples and branches<br>
in orchard environments using visual sensors. This network<br>
inherent challenges, sue<br>
enhances feature extraction capabilities through spatial<br>
omission, and poor in orchard environments using visual sensors. This network inherent challenges, such<br>enhances feature extraction capabilities through spatial omission, and poor segmentat<br>experimental results demonstrate that the optimal enhances feature extraction capabilities through spatial omission, and poor segmentation pyramid pooling and a gated feature pyramid structure. these advances have facilitated is experimental results demonstrate that the pyramid pooling and a gated feature pyramid structure. these advances have facili<br>Experimental results demonstrate that the optimal model<br>segmentation of specific tare<br>diversa segmentation accuracy of 87.6% and an F1 score Experimental results demonstrate that the optimal model<br>
segmentation of specific<br>
achieves a segmentation accuracy of 87.6% and an F1 score<br>
proposed new direction<br>
of 77.2%. Mo et al.[14] proposed a method for the seman achieves a segmentation accuracy of 87.6% and an F1 score<br>of 77.2%. Mo et al.[14] proposed a method for the semantic network architectures. The encoder utilizes a lightwight MobileNet facility environents present charmelit of 77.2%. Mo et al.[14] proposed a method for the semantic network architectures.<br>
segmentation of apples based on an improved DeepLabV3+ However, the complex<br>
architecture. The encoder utilizes a lightweight MobileNet fac segmentation of apples based on an improved DeepLabV3+<br>
architecture. The encoder utilizes a lightweight MobileNet<br>
facility environments prese<br>
module for feature extraction and employs depthwise in lighting angles, occlu architecture. The encoder utilizes a lightweight MobileNet facility environments present channol<br>and module for feature extraction and employs depthwise in lighting angles, occlusion or<br>separable convolution instead of sta module for feature extraction and employs depthwise in lighting angles, occl<br>separable convolution instead of standard convolution. This angles, and the similar<br>imodel achieves a pixel accuracy (PA) of 87.1%. Semantic fir separable convolution instead of standard convolution. This angles, and the similar model achieves a pixel accuracy (PA) of 95.3% and a mean These factors impact site<br>intersection over union (MIoU) of 87.1%. Semantic furth model achieves a pixel accuracy (PA) of 95.3% and a mean These factors impact seg<br>intersection over union (MIoU) of 87.1%. Semantic further research and imp<br>egementation is also commonly used to segment rotten parts semant intersection over union (MIoU) of 87.1%. Semantic further research and in<br>segmentation is also commonly used to segment rotten parts<br>fruits. For instance, Matsui [15] trained and validated a improved Swin-Unet [<br>U-net++ mo segmentation is also commonly used to segment rotten parts<br>
commentic segmentation mood fruits. For instance, Matsui [15] trained and validated a<br>
improved Swin-Unet [22]<br>
U-net++ model on X-ray avocado images to detect in of fruits. For instance, Matsui [15] trained and validated a improved<br>
U-net++ model on X-ray avocado images to detect internal study inc<br>
fruit rot, achieving an accuracy of 98%. Roy [16] constructed (1) To<br>
a semantic se net++ model on X-ray avocado images to detect internal<br>in tot, achieving an accuracy of 98%. Roy [16] constructed (1) To more accurate<br>semantic segmentation model based on En-UNet to similar in color to the b<br>ignment rotte fruit rot, achieving an accuracy of 98%. Roy [16] constructed (1) To more accurately<br>a semantic segmentation model based on En-UNet to similar in color to the back<br>segment rotten parts in apple RGB images, achieving traini a semantic segmentation model based on En-UNet to similar in color to the based meant rotten parts in apple RGB images, achieving training Attention Gate (AG) nand validation accuracies of 97.46% and 97.54%, Attention coef segment rotten parts in apple RGB images, achieving training Attention Gate (AG) module vand validation accuracies of 97.46% and 97.54%. Attention coefficients are expectively. These studies highlight the significance of i

and validation accuracies of 97.46% and 97.54% Attention coefficients are respectively. These studies highlight the significance of importance of each feature, a image semantic segmentation technology in the agricultural f respectively. These studies highlight the significance of importance of each feature, allowing image semantic segmentation technology in the agricultural features associated with green domain, particularly in cases where t image semantic segmentation technology in the agricultural features associated with green<br>domain, particularly in cases where there is a pronounced irrelevant background regions.<br>color difference between fruits and the bac domain, particularly in cases where there is a pronounced irrelevant background regions.<br>
color difference between fruits and the background. Accurate image segmentation can improve agricultural automation features and opt color difference between fruits and the background. Accurate (2) To achieve multi-scale<br>
image segmentation can improve agricultural automation<br>
efficiency, enhance fruit quality assessment, and optimize deges, this study image segmentation can improve agricultural automation<br>
efficiency, enhance fruit quality assessment, and optimize<br>
edges, this study properace discussed effection and management strategies.<br>
In complex situations where t efficiency, enhance fruit quality assessment, and optimize edges, this study proposes the A<br>disease detection and management strategies.<br>
In complex situations where the target and the background keeping the parameter quan disease detection and management strategies. section, which enla<br>
In complex situations where the target and the background keeping the parame<br>
colors are similar, deep learning-based image semantic model's ability to ha<br> In complex situations where the target and the background<br>
colors are similar, deep learning-based image semantic model's ability to handle contex<br>
segmentation techniques face significant challenges, scales.<br>
primarily i colors are similar, deep learning-based image semantic model's ability to handle consentation techniques face significant challenges, scales.<br>
primarily in distinguishing between the target and the (3) Experiments conduct segmentation techniques face significant challenges, scales.<br>
primarily in distinguishing between the target and the  $(3)$  Experiments conductional. To address this issue, several studies have made dataset demonstrate that primarily in distinguishing between the target and the (3) Experiments conducted<br>background. To address this issue, several studies have made dataset demonstrate that this<br>ignificant progress. For example, Li [17] proposed background. To address this issue, several studies have made<br>
significant progress. For example, Li [17] proposed an state-of-the-art techniques in<br>
optimized U-Net model by integrating residual blocks and<br>
gated convolut significant progress. For example, Li [17] proposed an state-of-the-art technic<br>optimized U-Net model by integrating residual blocks and proving to be more suita<br>used Artous Spatial Pyramid Pooling (ASPP) to merge Edge<br>fe optimized U-Net model by integrating residual blocks and<br>
gated convolutions to develop the Edge structure. They also<br>
facility environments.<br>
used Atrous Spatial Pyramid Pooling (ASPP) to merge Edge<br>
features with the hig gated convolutions to develop the Edge structure. They also<br>
since Arrous Spatial Pyramid Pooling (ASPP) to merge Edge<br>
features with the high-level features of U-Net, significantly<br>
improving the segmentation accuracy fo used Atrous Spatial Pyramid Pooling (ASPP) to merge Edge<br>
features with the high-level features of U-Net, significantly<br>
improving the sogmentation accuracy for green apples and<br>
FHE [18] enhanced the DeepLabV3+ model by r features with the high-level features of U-Net, significantly<br>
improving the segmentation accuracy for green apples and<br>
enhanced the DeepLaby3+ model by replacing its<br>
He [18] enhanced the DeepLaby3+ model by replacing it improving the segmentation accuracy for green apples and<br>
enhancing the model's generalization ability. Subsequently,<br>
He [18] enhanced the DeepLabV3+ model by replacing its<br>
databone with MobileNetV2, introducing the Suff enhancing the model' s generalization ability. Subsequently,<br>
He [18] enhanced the DeepLabV3+ model by replacing its<br>
backbone with MobileNetV2, introducing the Shuffle<br>
Actention Mechanism, and replacing the activation fi He [18] enhanced the DeepLabV3+ model by replacing its<br>
backbone with MobileNetV2, introducing the Shuffle<br>
Attention Mechanism, and replacing the activation function<br>
attention facility agricultu<br>
Meta-ACONC. This enhanc backbone with MobileNetV2, introducing the Shuffle<br>
Attention Mechanism, and replacing the activation function<br>
with Meta-ACONC. This enhancement increased the MIoU<br>
leaves complicates recom<br>
metric for green banana crown Attention Mechanism, and replacing the activation function<br>
with Meta-ACONC. This enhancement increased the MIoU<br>
metric for green banana crown segmentation to 85.75% and<br>
metric for green banana crown segmentation to 85.7 with Meta-ACONC. This enhancement increased the MIoU<br>metric for green banana crown segmentation to 85.75% and<br>the MPA to 91.41%. Yan [19] proposed a lightweight<br>the MPA to 91.41%. Yan [19] proposed a lightweight<br>convoluti

g Letters<br>
among green leaves using hyperspectral inputs,<br>
demonstrating the effectiveness of this method in generating<br>
more stable and less noisy segmentation results [21].<br>
Although deep learning has made significant pr g Letters<br>among green leaves using hyperspectral inputs,<br>demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant progre g Letters<br>among green leaves using hyperspectral inputs,<br>demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant progre g Letters<br>
among green leaves using hyperspectral inputs,<br>
demonstrating the effectiveness of this method in generating<br>
more stable and less noisy segmentation results [21].<br>
Although deep learning has made significant pr **g Letters**<br>among green leaves using hyperspectral inputs,<br>demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant prog **g Letters**<br>among green leaves using hyperspectral inputs,<br>demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant prog **g Letters**<br>among green leaves using hyperspectral inputs,<br>demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant prog **g Letters**<br>among green leaves using hyperspectral inputs,<br>demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant prog **g Letters**<br>among green leaves using hyperspectral inputs,<br>demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant prog **Exercise 15 Secure 15 Secure 15 Secure 15 Secure 16 Secure 16** among green leaves using hyperspectral inputs,<br>demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant progress in<br>imag among green leaves using hyperspectrem<br>demonstrating the effectiveness of this method<br>more stable and less noisy segmentation<br>Although deep learning has made significan<br>image segmentation, applications involving sim<br>target nong green leaves using hyperspectral inputs,<br>monstrating the effectiveness of this method in generating<br>ore stable and less noisy segmentation results [21].<br>though deep learning has made significant progress in<br>age segmen demonstrating the effectiveness of this method in generating<br>more stable and less noisy segmentation results [21].<br>Although deep learning has made significant progress in<br>image segmentation, applications involving similarl more stable and less noisy segmentation results [21].<br>Although deep learning has made significant progress in<br>image segmentation, applications involving similarly colored<br>targets and backgrounds in specific scenarios conti Although deep learning has made significant progress in<br>image segmentation, applications involving similarly colored<br>targets and backgrounds in specific scenarios continue to face<br>inherent challenges, such as target miscla

image segmentation, applications involving similarly colored<br>targets and backgrounds in specific scenarios continue to face<br>inherent challenges, such as target misclassification,<br>omission, and poor segmentation of fruit ed targets and backgrounds in specific scenarios continue to face<br>inherent challenges, such as target misclassification,<br>omission, and poor segmentation of fruit edges. Nonetheless,<br>these advances have facilitated research in inherent challenges, such as target misclassification,<br>omission, and poor segmentation of fruit edges. Nonetheless,<br>these advances have facilitated research into the semantic<br>segmentation of specific targets, such as green omission, and poor segmentation of fruit edges. Nonetheless,<br>these advances have facilitated research into the semantic<br>segmentation of specific targets, such as green tomatoes, and<br>proposed new directions for designing an these advances have facilitated research in<br>segmentation of specific targets, such as gree<br>proposed new directions for designing and<br>network architectures.<br>However, the complexity and unstruct<br>facility environments present gmentation of specific targets, such as green tomatoes, and<br>oposed new directions for designing and optimizing deep<br>twork architectures.<br>However, the complexity and unstructured nature of<br>cility environments present challe proposed new directions for designing and optimizing deep<br>network architectures.<br>However, the complexity and unstructured nature of<br>facility environments present challenges, including variations<br>in lighting angles, occlusi network architectures.<br>
However, the complexity and unstructured nature of<br>
facility environments present challenges, including variations<br>
in lighting angles, occlusion or overlap of fruits, collection<br>
angles, and the si However, the complexity and unstructured nature of facility environments present challenges, including variations in lighting angles, occlusion or overlap of fruits, collection angles, and the similarity of green fruits to facility environments present challenges, including variations<br>in lighting angles, occlusion or overlap of fruits, collection<br>angles, and the similarity of green fruits to the background.<br>These factors impact segmentation in lighting angles, occlusion or overlap of fruits, collection<br>angles, and the similarity of green fruits to the background.<br>These factors impact segmentation accuracy and necessitate<br>further research and improvements. Thi

angles, and the similarity of green fruits to the background.<br>These factors impact segmentation accuracy and necessitate<br>further research and improvements. This paper proposes a<br>semantic segmentation model for green tomato ese factors impact segmentation accuracy and necessitate<br>ther research and improvements. This paper proposes a<br>mantic segmentation model for green tomatoes based on an<br>proved Swin-Unet [22]. The main contributions of this<br> further research and improvements. This paper proposes a<br>semantic segmentation model for green tomatoes based on an<br>improved Swin-Unet [22]. The main contributions of this<br>study include:<br>(1) To more accurately segment gree semantic segmentation model for green tomatoes based on an<br>improved Swin-Unet [22]. The main contributions of this<br>study include:<br>(1) To more accurately segment green tomatoes, which are<br>similar in color to the background,

improved Swin-Unet [22]. The main contributions of this<br>study include:<br>(1) To more accurately segment green tomatoes, which are<br>similar in color to the background, this study incorporates the<br>Attention Gate (AG) module wit study include:<br>
(1) To more accurately segment green tomatoes, which are<br>
similar in color to the background, this study incorporates the<br>
Attention Gate (AG) module within the skip connections.<br>
Attention coefficients are (1) To more accurately segment green tomatoes, which are similar in color to the background, this study incorporates the Attention Gate (AG) module within the skip connections. Attention coefficients are designed to evalua scales. tention Gate (AG) module within the skip connections.<br>
tention coefficients are designed to evaluate the<br>
portance of each feature, allowing the model to focus on<br>
tures associated with green tomatoes while suppressing<br>
el Attention coefficients are designed to evaluate the<br>importance of each feature, allowing the model to focus on<br>features associated with green tomatoes while suppressing<br>irrelevant background regions.<br>(2) To achieve multi-s importance of each feature, allowing the model to focus on<br>features associated with green tomatoes while suppressing<br>irrelevant background regions.<br>(2) To achieve multi-scale extraction of green tomato<br>features and optimiz features associated with green tomatoes while suppressing<br>irrelevant background regions.<br>(2) To achieve multi-scale extraction of green tomato<br>features and optimize the smoothness of their segmentation<br>edges, this study pr irrelevant background regions.<br>
(2) To achieve multi-scale extraction of<br>
features and optimize the smoothness of their<br>
edges, this study proposes the ASPP module in<br>
section, which enlarges the feature receptiv<br>
keeping ptimize the smoothness of their segmentation<br>dy proposes the ASPP module in the bottleneck<br>in enlarges the feature receptive field while<br>arameter quantity unchanged, enhancing the<br>to handle contextual information at differ

Equality proposes the ASFF module in the bottlenck<br>
section, which enlarges the feature receptive field while<br>
keeping the parameter quantity unchanged, enhancing the<br>
model's ability to handle contextual information at di From the parameter quantity unchanged, enhancing the parameter quantity unchanged, enhancing the dodel's ability to handle contextual information at different alles.<br>
(3) Experiments conducted on a custom green tomato tase segmentation in factorization of the resulting the product of shility to handle contextual information at different scales.<br>
(3) Experiments conducted on a custom green tomato dataset demonstrate that this method outperfor scales.<br>
Scales.<br>
(3) Experiments conducted on a custom green tomato<br>
dataset demonstrate that this method outperforms other<br>
state-of-the-art techniques in accuracy and robustness.<br>
proving to be more suitable for segment (3) Experiments conducted on a custom green tomato<br>dataset demonstrate that this method outperforms other<br>state-of-the-art techniques in accuracy and robustness.<br>proving to be more suitable for segmenting green tomatoes in

dataset demonstrate that this method outperforms other<br>dataset demonstrate that this method outperforms other<br>state-of-the-art techniques in accuracy and robustness.<br>proving to be more suitable for segmenting green tomatoe *Image Acquisition* Experimental Exerces Compare the-of-the-art techniques in accuracy and r<br> *NATERIALS AND METHODS*<br> *II.* MATERIALS AND METHODS<br> *Green Tomato Fruit Dataset*<br>
This study aims to address the challenge of Image Theorem and Contains and Theorem and Solar Theorem and Solar University the more suitable for segmenting green tomatoes in Sility environments.<br>
II. MATERIALS AND METHODS<br> *Green Tomato Fruit Dataset*<br>
This study aim Froving to be introducted on organisating given tends<br>to facility environments.<br>II. MATERIALS AND METHODS<br>A. Green Tomato Fruit Dataset<br>This study aims to address the challenge of green tomato<br>segmentation in facility agri II. MATERIALS AND METHODS<br> *A.* Green Tomato Fruit Dataset<br>
This study aims to address the challenge of g<br>
segmentation in facility agricultural environment<br>
color similarity between green tomato fruits and<br>
leaves compli II. MATERIALS AND METHODS<br> *Green Tomato Fruit Dataset*<br>
This study aims to address the challenge of green tomato<br>
gmentation in facility agricultural environments, where the<br>
lor similarity between green tomato fruits an A. Green Tomato Fruit Dataset<br>
This study aims to address the challenge of green tomato<br>
segmentation in facility agricultural environments, where the<br>
color similarity between green tomato fruits and background<br>
leaves c A. Green Tomato Fruit Dataset<br>This study aims to address the challenge of green tomato<br>segmentation in facility agricultural environments, where the<br>color similarity between green tomato fruits and background<br>leaves compl This study aims to address the challenge of green tomato<br>gmentation in facility agricultural environments, where the<br>lor similarity between green tomato fruits and background<br>wes complicates recognition, often resulting in segmentation in facility agricultural environments, where the<br>color similarity between green tomato fruits and background<br>leaves complicates recognition, often resulting in fruit<br>omission or confusion with branches and lea

color similarity between green tomato fruits and background<br>leaves complicates recognition, often resulting in fruit<br>omission or confusion with branches and leaves.<br>*1) Image Collection*<br>Image Acquisition Location: Greenho leaves complicates recognition, often resulting in fruit omission or confusion with branches and leaves.<br> *1) Image Collection*<br>
Image Acquisition Location: Greenhouse, Hetong Village,<br>
Shangkou Town, Shouguang City, Weifa



















g Distant view green tomato image<br>g Distant view green tomato image<br>Fig.1 Images of green tomato finits in different environments<br>The captured images are presented in Fig. 1. Figures 1a to to 640 × 640 pixels. The<br>Id displ g Distant view green tomato image<br>
g Distant view green tomato image<br>
Fig.1 Images of green tomato fruits in different environments<br>
The captured images are presented in Fig. 1. Figures 1a to to 640 × 640 pixels. This ad<br> g Distant view green tomato image<br>Fig.1 Images of green tomato fruits in different environments<br>Fig.1. Figures 1 at to to 640 × 640 pixels. This<br>Id display green tomatoes under different lighting conditions, green tomatoe g Distant view green tomato image<br>
19 Eig.1 Images of green tomato finits in different environments<br>
The captured images are presented in Fig. 1. Figures 1a to to 640 × 640 pixels. This<br>
1d display green tomatoes under dif provides and viewpoints, simulating perspectives typical simulation of the september of the separation of the separation of the captured images are presented in Fig. 1. Figures 1a to to 640 × 640 pixels. This is a ld displ Fig.1 Images of green tomato image<br>
Fig.1 Images of green tomato image<br>
Fig.1 Images of green tomato fruits in different environments<br>
The captured images are presented in Fig. 1. Figures 1a to to 640 × 640 pixels. This<br>
1 g Distant view green tomato image<br>
Fig.1 Images of green tomato fruits in different environments<br>
The captured images are presented in Fig. 1. Figures 1a to to 640 × 640 pixels. This adji<br>
1d display green tomatoes under d g Distant view green tomato image<br>
Fig.1 Images of green tomato fruits in different environments<br>
The captured images are presented in Fig.1. Figures 1a to to 640 × 640 pixels. This adj<br>
1d display green tomatoes under dif Fig.1 Images of green tomato fruits in different environments<br>
The captured images are presented in Fig. 1. Figures 1a to to 640 × 640 pixels. This<br>
1d display green tomatoes under different lighting conditions, green tom Fig.1 Images of green tomato fruits in different environm<br>
The captured images are presented in Fig.<br>
1d display green tomatoes under different light<br>
including natural daytime illumination (both<br>
lighting) and nighttime i The captured images are presented in Fig. 1. Figu display green tomatoes under different lighting coluding natural daytime illumination (both front a hting) and nighttime illumination by LED lights to 1g illustrate tomato The captured images are presented in Fig. 1. Figures 1a to to 640  $\times$  640 pixels.<br>
display green tomatoes under different lighting conditions, green tomatoes egn<br>
cluding natural daytime illumination (both front and back 1d display green tomatoes under different lighting conditions, green tomato segmentation<br>
including natural daytime illumination by LED lights. Figures adaptable to the segmentation<br>
lighting) and nighttime illumination by including natural daytime illumination (both front and back<br>
adaptable to the seg<br>
lighting) and nighttime illumination by LED lights. Figures Existing datasets f<br>
le to 1g illustrate tomato images taken from various<br>
prox lighting) and nighttime illumination by LED lights. Figures Existing datasets for green ton le to 1g illustrate tomato images taken from various primarily designed for classifice proximities and viewpoints, simulating pers

1e to 1g illustrate tomato images taken from various<br>
primarily designed for classificat<br>
proximities and viewpoints, simulating perspectives typical<br>
one cassary for semantic segment<br>
of picking equipment in a real orchar movimities and viewpoints, simulating perspectives typical<br>of picking equipment in a real orchard environment. Figures the LabelMe [23] software w<br>th and 1i show examples of significant shading and annotate the datasets in of picking equipment in a real orchard environment. Figures the LabelMe [23] softwas the and 11 show examples of significant shading and annotate the datasets in overlapping in facility-based agricultural environments, wit The and 1 is show examples of significant shading and annotate the datasets<br>overlapping in facility-based agricultural environments, with<br>fruits obscuring one another and branches and leaves causing image, thereby provi<br>in

Final to the segmentation of low-resolution images.<br>
The segmentation of low-resolution images.<br>
Existing datasets for green tomato image classification are<br>
primarily designed for classification tasks and lack the labels<br> Existing datasets for green tomato image<br>to 640 × 640 pixels. This adjustment aims to optimize the<br>green tomato segmentation network, making it more<br>daptable to the segmentation of low-resolution images.<br>Existing datasets tomato image<br>
i Overlapping green tomato image<br>
to 640 × 640 pixels. This adjustment aims to optimize the<br>
green tomato segmentation network, making it more<br>
adaptable to the segmentation of low-resolution images.<br>
Existin necessary for semantic segmentation. To address involved and the User Captain and the segmentation of low-resolution images.<br>Existing datasets for green tomato image classification are primarily designed for classification tomato image<br>
to 640 × 640 pixels. This adjustment aims to optimize the<br>
green tomato segmentation network, making it more<br>
adaptable to the segmentation of low-resolution images.<br>
Existing datasets for green tomato image tomato image in Overlapping green tomato image<br>to 640  $\times$  640 pixels. This adjustment aims to optimize the<br>green tomato segmentation network, making it more<br>adaptable to the segmentation of low-resolution images.<br>Existin tomato image i Overlapping green tomato image<br>to 640 × 640 pixels. This adjustment aims to optimize the<br>green tomato segmentation network, making it more<br>adaptable to the segmentation of low-resolution images.<br>Existing dat image, the station minimism and the station of low-resolution in the segmentation of low-resolution images.<br>Existing datasets for green tomato image classification are primarily designed for classification tasks and lack t to 640  $\times$  640 pixels. This adjustment aims to optimize the green tomato segmentation network, making it more adaptable to the segmentation of low-resolution images. Existing datasets for green tomato image classificatio to 640  $\times$  640 pixels. This adjustment aims to optimize the green tomato segmentation network, making it more adaptable to the segmentation of low-resolution images. Existing datasets for green tomato image classificatio to 640  $\times$  640 pixels. This adjustment aims to optimize the<br>green tomato segmentation network, making it more<br>adaptable to the segmentation of low-resolution images.<br>Existing datasets for green tomato image classificatio green tomato segmentation network, making it more<br>adaptable to the segmentation of low-resolution images.<br>Existing datasets for green tomato image classification are<br>primarily designed for classification tasks and lack the adaptable to the segmentation of low-resolution images.<br>Existing datasets for green tomato image classification are<br>primarily designed for classification tasks and lack the labels<br>necessary for semantic segmentation. To ad xisting datasets for green tomato image classification are<br>imarily designed for classification tasks and lack the labels<br>cessary for semantic segmentation. To address this issue,<br>elections at a subset of the datasets in de primarily designed for classification tasks and lack the labels<br>necessary for semantic segmentation. To address this issue,<br>the LabelMe [23] software was employed to manually<br>annotate the datasets in detail. This process i necessary for semantic segmentation. To address this issue,<br>the LabelMe [23] software was employed to manually<br>annotate the datasets in detail. This process involved<br>generating category labels and annotation points for eac the LabelMe [23] software was employed to manually<br>annotate the datasets in detail. This process involved<br>generating category labels and annotation points for each<br>image, thereby providing the essentation ground truth<br>info annotate the datasets in detail. This process involved<br>generating category labels and annotation points for each<br>image, thereby providing the essentation ground truth<br>information for semantic segmentation. All annotation d





Dataset Class Images Instances Small Medium<br>  $(0 \leq \text{area} < 32^2)$   $(32^2 \leq \text{area} < 96^2)$ <br>
Train tomato 1066 4752 3  $(0.06\%)$  232(4.88%)<br>
Test tomato 267 1442 5  $(0.35\%)$  187(12.97%)<br>
Green tomato fruits were categorize  $(0 \le \text{area} < 32^2)$   $(32^2 < \text{area} \le 12^2)$ <br>
Train tomato 1066 4752 3 (0.06%) 232(4.8<br>
Test tomato 267 1442 5 (0.35%) 187(12.9<br>
Green tomato fruits were categorized into small-scale, Swin Transformer wit<br>
medium-scale, a  $(0 \leq \text{area} < 32^2)$   $(32 \leq \text{area} < 96^2)$ <br>
Train tomato 1066 4752 3 (0.06%) 232(4.88%)<br>
Test tomato 267 1442 5 (0.35%) 187(12.97%)<br>
Green tomato fruits were categorized into small-scale, Swin Transformer with the<br>
medi Train tomato 1066 4752 3  $(0.06\%)$  232(4.88%)<br>
Test tomato 267 1442 5  $(0.35\%)$  187(12.97%)<br>
Green tomato fruits were categorized into small-scale, Swin Transformer with the<br>
medium-scale, and large-scale classes accordin Test tomato  $1000$   $-1/2$ <br>Test tomato  $267$   $144$ <br>Green tomato fruits were categorized into small<br>medium-scale, and large-scale classes according<br>criteria used in the Microsoft COCO [24] dataset. The<br>of each fruit instanc First tomato 267 1442 5 (0.53%) 167<br>
Green tomato fruits were categorized into small-scale, Swin Transformer<br>
medium-scale, and large-scale classes according to the capabilities of U-N<br>
criteria used in the Microsoft COCO Green tomato fruits were categorized into small-scale,<br>
dium-scale, and large-scale classes according to the capabilities of U-Net, resulteria used in the Microsoft COCO [24] dataset. The area<br>
teria used in the Microsoft medium-scale, and large-scale classes according to the<br>
criteria used in the Microsoft COCO [24] dataset. The area<br>
of each fruit instance was determined by the number of<br>
pixels in its corresponding mask to evaluate the p oriential time intervention, overlapping branches and foliage, and variable signify that the substitution of the substitution of the substitution, with the substitution, with is controlled to the algorithm. Table I provide

pixels in its corresponding mask to evaluate the performance<br>of the algorithm. Table I provides the relevant details. It is images with complex back<br>noteworthy that the number of small-scale fruits is relatively consists For the algorithm. Table I provides the relevant details. It is<br>
images with complex back<br>
noteworthy that the number of small-scale fruits is relatively<br>
noteworthy that the number of small-scale fruits is relatively<br>
con From the similar to the background, their boundaries are often information and Skip Connection and Skip Connection and Skip Connection are counting for 0.35%.<br>
In facility-based agricultural environments, image and downsam Ion, with only three small targets in the training set,<br>
Ion, with only three small targets in the validation set<br>
accounting for 0.35%.<br>
Einear Embedding technication of Swin-Unet Segmentation Model<br>
In facility-based agr representing 0.06%, and small targets in the validation set<br>
representing 0.06%, and small targets in the validation set<br>
accounting for 0.35%.<br> *B. Optimization of Swin-Unet Segmentation Model*<br>
In facility-based agricult accounting for 0.35%.<br> **Examples are the interpretation Model**<br>
In facility-based agricultural environments, image<br>
In facility-based agricultural environments, image<br>
and downsampling the incorporates and Artous<br>
acquisi **Examing the set of**<br> *B. Optimization of Swin-Unet Segmentation Model*<br>
and downsampling through<br>
and downsampling through<br>
accupation is challenging due to complex backgrounds<br>
acculusion, overlapping branches and folia *B. Optimization of Swin-Unet Segmentation Model*<br>
In facility-based agricultural environments, image<br>
acquisition is challenging due to complex backgrounds,<br>
ighting condusion, overlapping branches and doinge, and variab In facility-based agricultural environments, image<br>acquisition is challenging due to complex backgrounds,<br>acquisition, overlapping branches and foliage, and variable<br>including downlight, backlight, and foliage and expand acquisition is challenging due to complex backgrounds,<br>
including coclusion, overlapping branches and foliage, and variable<br>
ilghting conditions (including downlight, backlight, and<br>
ilghting conditions (including downligh occlusion, overlapping branches and foliage, and variable<br>
ighting conditions (including downlight, backlight, and<br>
ighttime environments), all of which can adversely affect<br>
image quality. Specifically, for green fruits w lighting conditions (including downlight, backlight, and<br>
inghttime environments), all of which can adversely affect<br>
image quality. Specifically, for green fruits with colors<br>
similar to the background, their boundaries mighttime environments), all of which can adversely affect layers, a<br>image quality. Specifically, for green fruits with colors image sizual<br>similar to the background, their boundaries are often informal<br>indistinct, which image quality. Specifically, for green fruits with colors<br>
implairs implaires. Additionally, an<br>
imilar to the background, their boundaries are often<br>
indistinct, which significantly complicates accurate<br>
information part imilar to the background, their boundaries are often<br>
indistinct, which significantly complicates accurate<br>
information while suppress<br>
segmentation [25]. Additionally, a notable issue in the<br>
information while suppress<br>
i indistinct, which significantly complicates accurate<br>
segmentation Valie segmentation (25). Additionally, a notable issue in the<br>
architecture information while<br>
directively integrates<br>
fruits. The lack of adequate labeled segmentation [25]. Additionally, a notable issue in the improved segmentation accomparison of specific types of fruits, such as green fruits. The lack of adequate labeled samples to train issues in fruits. The lack of adeq

of each fruit instance was determined by the number of<br>pixels in its corresponding mask to evaluate the performance<br>information, which is crucial<br>of the algorithm. Table I provides the relevant details. It is<br>mages with co Target Amount<br>
Small Medium Large<br>
area<32<sup>2</sup>)  $(32^2 \le \text{area} < 96^2)$   $(96^2 \le \text{area})$ <br>  $0.06\%$   $232(4.88\%)$   $4517(95.06\%)$ <br>  $(0.35\%)$   $187(12.97\%)$   $1250(86.68\%)$ <br>
Swin Transformer with the high-precision segmentation<br>
c Small Medium Large<br>
area<32<sup>2</sup>)  $(32^2 \le \text{area} \le 96^2)$   $(96^2 \le \text{area})$ <br>  $0.06\%$   $232(4.88\%)$   $4517(95.06\%)$ <br>  $(0.35\%)$   $187(12.97\%)$   $1250(86.68\%)$ <br>
Swin Transformer with the high-precision segmentation<br>
capabilities of U consists of four main components: Encoder, and Skip Consists of four main components: Encoder, and Skip Consists of the main compon area<32<sup>2</sup>)  $(32\textdegree\textdegree\cdot486\textdegree\textdegree)$   $(96\textdegree\textdegree\cdot486\textdegree\textdegree)$ <br>  $0.06\%$   $232(4.88\%)$   $4517(95.06\%)$ <br>  $(0.35\%)$   $187(12.97\%)$   $1250(86.68\%)$ <br>
Swin Transformer with the high-precision segmentation<br>
capabilities of U-Net, resu 0.06%)  $232(4.88\%)$   $4517(95.06\%)$ <br>  $(0.35\%)$   $187(12.97\%)$   $1250(86.68\%)$ <br>
Swin Transformer with the high-precision segmentation<br>
capabilities of U-Net, resulting in an imovative framework.<br>
Through its unique design, th (0.35%) 187(12.97%) 1250(86.68%)<br>
Swin Transformer with the high-precision segmentation<br>
capabilities of U-Net, resulting in an imovative framework.<br>
Through its unique design, the model effectively addresses<br>
long-range (0.35%) 187(12.97%) 1250(86.68%)<br>
Swin Transformer with the high-precision segmentation<br>
capabilities of U-Net, resulting in an innovative framework.<br>
Through its unique design, the model effectively addresses<br>
long-range Swin Transformer with the high-precision segmentation<br>capabilities of U-Net, resulting in an innovative framework.<br>Through its unique design, the model effectively addresses<br>long-range dependency issues while preserving sp Swin Transformer with the high-precision segmentation capabilities of U-Net, resulting in an innovative framework.<br>Through its unique design, the model effectively addresses long-range dependency issues while preserving sp capabilities of U-Net, resulting in an innovative framework.<br>Through its unique design, the model effectively addresses<br>long-range dependency issues while preserving spatial<br>information, which is crucial for segmenting gre Through its unique design, the model effectively addresses<br>long-range dependency issues while preserving spatial<br>information, which is crucial for segmenting green tomato<br>images with complex backgrounds. The optimized mode long-range dependency issues while preserving spatial<br>information, which is crucial for segmenting green tomato<br>images with complex backgrounds. The optimized model<br>consists of four main components: Encoder, Bottleneck,<br>De information, which is crucial for segmenting green tomato images with complex backgrounds. The optimized model<br>consists of four main components: Encoder, Bottleneck,<br>Decoder, and Skip Connection (Fig. 3). In the Encoder st images with complex backgrounds. The optimized model<br>consists of four main components: Encoder, Bottleneck,<br>Decoder, and Skip Connection (Fig. 3). In the Encoder stage,<br>the model adjusts channel numbers using Patch Partiti consists of four main components: Encoder, Bottleneck,<br>Decoder, and Skip Connection (Fig. 3). In the Encoder stage,<br>the model adjusts channel numbers using Patch Partition and<br>Linear Embedding techniques to achieve feature Decoder, and Skip Connection (Fig. 3). In the Encoder stage,<br>the model adjusts channel numbers using Patch Partition and<br>Linear Embedding techniques to achieve feature extraction<br>and downsampling through multiple Swin Tran the model adjusts channel numbers using Patch Partition and<br>Linear Embedding techniques to achieve feature extraction<br>and downsampling through multiple Swin Transformer<br>Blocks and Patch Merging layers. The Bottleneck stage Linear Embedding techniques to achieve feature extraction<br>and downsampling through multiple Swin Transformer<br>Blocks and Patch Merging layers. The Bottleneck stage<br>incorporates an Atrous Spatial Pyramid Pooling (ASPP)<br>modul and downsampling through multiple Swin Transformer<br>Blocks and Patch Merging layers. The Bottleneck stage<br>incorporates an Atrous Spatial Pyramid Pooling (ASPP)<br>module [26] to capture image information at various scales<br>and ocks and Patch Merging layers. The Bottleneck stage<br>corporates an Atrous Spatial Pyramid Pooling (ASPP)<br>Jodule [26] to capture image information at various scales<br>d expand receptive fields. In the Decoder stage, multiple<br>v corporates an Atrous Spatial Pyramid Pooling (ASPP)<br>
odule [26] to capture image information at various scales<br>
d expand receptive fields. In the Decoder stage, multiple<br>
vin Transformer modules, along with Patch Expanding module [26] to capture image information at various scales<br>and expand receptive fields. In the Decoder stage, multiple<br>Swin Transformer modules, along with Patch Expanding<br>layers, are employed for upsampling and restoring and expand receptive fields. In the Decoder stage, multiple<br>Swin Transformer modules, along with Patch Expanding<br>layers, are employed for upsampling and restoring feature<br>map sizes. Additionally, an Attention Gate (AG) mod Swin Transformer modules, along with Patch Expanding<br>layers, are employed for upsampling and restoring feature<br>map sizes. Additionally, an Attention Gate (AG) module [27]<br>is introduced in skip connections to enhance target

layers, are employed for upsampling and restoring feature<br>map sizes. Additionally, an Attention Gate (AG) module [27]<br>is introduced in skip connections to enhance target feature<br>information while suppressing irrelevant det map sizes. Additionally, an Attention Gate (AG) module [27]<br>is introduced in skip connections to enhance target feature<br>information while suppressing irrelevant details for<br>improved segmentation accuracy. This entire proce is introduced in skip connections to enhance target feature<br>information while suppressing irrelevant details for<br>improved segmentation accuracy. This entire process<br>effectively integrates multi-scale information, enhancing information while suppressing irrelevant details for<br>improved segmentation accuracy. This entire process<br>effectively integrates multi-scale information, enhancing the<br>segmentation results of green tomato images.<br>I) Target improved segmentation accuracy. This entire process<br>effectively integrates multi-scale information, enhancing the<br>segmentation results of green tomato images.<br>*I) Target Feature Enhancement Module*<br>In complex scenarios whe



My32xHy32x8C<br>
My32xHy32x8C<br>
Eig. 3. Structure of the green tomato segmentation model optimized based on Swin-Unet<br>
suppressing less important areas. By calculating attention<br>
coefficients for each feature, the model adjus W/32xH/32x8C<br>
Fig. 3. Structure of the green tomato segmentation model optimized based on Swin-Unet<br>
suppressing less important areas. By calculating attention These parts are then we<br>
coefficients for each feature, the m W/32xH/32x8C<br>
Fig. 3. Structure of the green tomato segmentation model optimized based on Swin-Unet<br>
suppressing less important areas. By calculating attention<br>
coefficients for each feature, the model adjusts the weights  $W/32xH/32x8C$ <br>Fig. 3. Structure of the green tomato segmentation model opt<br>suppressing less important areas. By calculating<br>coefficients for each feature, the model adjusts tl<br>of the feature maps, directing the network's Skip connections merge features from the encoder and<br>
Skip continuous experimentation model optimized based on Swin-Unet<br>
ppressing less important areas. By calculating attention These parts are then we<br>
efficients for ea Fig. 3. Structure of the green tomato segmentation model optimized based on Swin-Unet<br>suppressing less important areas. By calculating attention These parts are then vecefficients for each feature, the model adjusts the w suppressing less important areas. By calculating attention<br>coefficients for each feature, the model adjusts the weights<br>of the feature maps, directing the network's focus towards<br>the target area rather than the background suppressing less important areas. By calculating attention<br>coefficients for each feature, the model adjusts the weights<br>of the feature maps, directing the network's focus towards<br>the target are are then the background. Th

coefficients for each feature, the model adjusts the weights<br>of the feature maps, directing the network's focus towards<br>the target area rather than the background. This optimization<br>the Sigmoid activation func<br>not only im of the feature maps, directing the network's focus towards<br>the target area rather than the background. This optimization<br>not only improves the model's learning process, making it<br>more efficient in extracting key features, the target area rather than the background. This optimization<br>
not only improves the model's learning process, making it<br>
significantly enhances overall accuracy. Specifically, in the segmentation of green tomatoes, it no not only improves the model's learning process, making it<br>
more efficient in extracting key features, but also weight in the fusion, wh<br>
significantly enhances overall accuracy. Specifically, in the<br>
significantly enhance more efficient in extracting key features, but also weight in the fusion, we<br>significantly enhances overall accuracy. Specifically, in the feature's weight is low<br>segmentation of green tomatoes, it notably increases alloc significantly enhances overall accuracy. Specifically, in the feature's weight is low<br>degeneration of green tomatoes, it notably increases allocation enables the m<br>diagnostic accuracy.<br>features based on the cord<br>decoder, segmentation of green tomatoes, it notably increases<br>
diagnostic accuracy.<br>
diagnostic accuracy.<br>
decoder, reserving the spatial integrity of the image while<br>
enhancing the model's ability to recognize details, such as<br>
e diagnostic accuracy.<br>
Skip connections merge features from the encoder and<br>
decoder, preserving the model's ability to recognize details, such as<br>
endancing the model's ability to recognize details, such as<br>
endages, whic Skip connections merge features from the encoder and<br>
decoder, preserving the spatial integrity of the image while<br>
enhancing the model's ability to recognize details, such as<br>
edges, which are crucial for the segmentatio decoder, preserving the spatial integrity of the image while<br>
enhancing the model's ability to recognize details, such as<br>
edges, which are crucial for the segmentation of green<br>
tomatoes. The incorporation of the Attenti enhancing the model's ability to recognize details, such as<br>
edges, which are crucial for the segmentation of green<br>
subsequently, through another<br>
tomatoes. The incorporation of the Attention Gate (AG)<br>
and the attention edges, which are crucial for the segmentation of green<br>
tomatoes. The incorporation of the Attention Gate (AG)<br>
module into skip connections further amplifies this<br>
are attention coefficier<br>
are transformed into skip come tomatoes. The incorporation of the Attention Gate (AG)<br>
module into skip connections further amplifies this<br>
advantage by enabling dynamic feature weight allocation.<br>
This allows the model to flexibly adjust its focus on transformations.

Swin Transformer<br>
Block x1<br>
Club and method and merged, and an<br>
antention coefficient  $\zeta$ , ranging from 0 to 1, is obtained via<br> Swin Transformer<br>Block x1<br>in-Unet<br>These parts are then weighted and merged, and an<br>attention coefficient  $\zeta$ , ranging from 0 to 1, is obtained via<br>the Sigmoid activation function by adding  $g_1$  and  $x_1$ . When<br> $\zeta$  is **For all of the COLOGET**<br> **FEATURE SET IS SET ASSESS TO THE SET AND A THE SET AND WEIGHT WEIGHT IN SURVEY IT IS CONSERVED WEIGHT WEIGHT WEIGHT WEIGHT I EXECUTE:**<br> **EXECUTE:**<br> **EXEC** Finction<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet<br>fin-Unet on the content of the image, and x<sub>1</sub>. When<br> $\zeta$  is close to 1, the correspo These parts are then weighted and merged, and an<br>attention coefficient  $\zeta$ , ranging from 0 to 1, is obtained via<br>the Sigmoid activation function by adding  $g_1$  and  $x_1$ . When<br> $\zeta$  is close to 1, the corresponding fea These parts are then weighted and merged, and an attention coefficient  $\zeta$ , ranging from 0 to 1, is obtained via the Sigmoid activation function by adding g<sub>1</sub> and x<sub>1</sub>. When  $\zeta$  is close to 1, the corresponding featu the Sigmoid activation function by adding  $g_1$  and  $x_1$ . When  $\zeta$  is close to 1, the corresponding feature has a higher weight in the fusion, whereas when  $\zeta$  is close to 0, the feature's weight is lower. This dynam  $\zeta$  is close to 1, the corresponding feature has a higher<br>weight in the fusion, whereas when  $\zeta$  is close to 0, the<br>feature's weight is lower. This dynamic feature weight<br>allocation enables the model to flexibly adjus

$$
\zeta = Sigmoid(g_1 + x_1) \tag{1}
$$

weight in the fusion, whereas when  $\zeta$  is close to 0, the<br>feature's weight is lower. This dynamic feature weight<br>allocation enables the model to flexibly adjust its focus on<br>features based on the content of the image, t there is a close to 0, the<br>
ther. This dynamic feature weight<br>
del to flexibly adjust its focus on<br>
ent of the image, thus determining<br>
ture.<br>  $l(g_1 + x_1)$  (1)<br>
another linear transformation  $\varphi$ ,<br>
s are adjusted to match weight in the tusion, whereas when  $\zeta$  is close to<br>feature's weight is lower. This dynamic feature<br>allocation enables the model to flexibly adjust its f<br>features based on the content of the image, thus dete<br>the importan features based on the content of the image, thus determining<br>the importance of each feature.<br> $\zeta = Sigmoid(g_1 + x_1)$  (1)<br>Subsequently, through another linear transformation  $\varphi$ ,<br>the attention coefficients are adjusted to match the importance of each feature.<br>  $\zeta = Sigmoid(g_1 + x_1)$  (1)<br>
Subsequently, through another linear transformation  $\varphi$ ,<br>
the attention coefficients are adjusted to match the<br>
dimensions of the feature map x. These coefficients

$$
x^{\sim} = \varphi(\zeta) \odot x \tag{2}
$$

 $\zeta$  = *Sigmoid* ( $g_1 + x_1$ ) (1)<br>
Subsequently, through another linear transformation  $\varphi$ ,<br>
the attention coefficients are adjusted to match the<br>
dimensions of the feature map x. These coefficients are then<br>
multiplied Subsequently, through another linear transformation  $\varphi$ ,<br>the attention coefficients are adjusted to match the<br>dimensions of the feature map x. These coefficients are then<br>multiplied by the encoder feature map x to obtai Subsequently, through another linear transformation  $\varphi$ ,<br>the attention coefficients are adjusted to match the<br>dimensions of the feature map x. These coefficients are then<br>multiplied by the encoder feature map x to obtai the attention coefficients are adjusted to match the<br>dimensions of the feature map x. These coefficients are then<br>multiplied by the encoder feature map x to obtain the<br>weighted feature map x<sup>-</sup>, thereby accomplishing feat dimensions of the feature map x. These coefficients are then<br>multiplied by the encoder feature map x to obtain the<br>weighted feature map x<sup>-</sup>, thereby accomplishing feature<br>selection and enhancement.<br> $x^{\sim} = \varphi(\zeta) \odot x$  (2



Fig.5. Atrous spatial pyramid pooling Module<br>
Prom 1x1 meanuple<br>
Prince Edge Restoration Module<br>
2) Multiscale Edge Restoration Module<br>
In the encoder-decoder architecture used for green tomato (ASPP) module is illustrate Fig.5. Atrous spatial pyramid pooling Module<br>
2) Multiscale Edge Restoration Module<br>
11 the encoder-decoder architecture used for green tomato (ASPP) module is empimage segmentation, the encoder expands the receptive fiel Aspect of the mechanism for the product of the mechanism of the mechanism and the mechanism for the encoder architecture used for green tomato (ASPP) module is employed are segmentation, the encoder expands the receptive Fig.5. Atrous spatial pyramid pooling Module<br>
2) Multiscale Edge Restoration Module<br>
In the encoder-decoder architecture used for green tomato (ASPP) module is employed.<br>
image segmentation, the encoder expands the recepti Fig.5. Atrous spatial pyramid pooling Module<br>
2) Multiscale Edge Restoration Module<br>
In the encoder-decoder architecture used for green tomato (ASPP) module is employed<br>
image segmentation, the encoder expands the recepti

2) Multiscale Edge Restoration Module<br>
In the encoder architecture used for green tomato (ASPP) module is image segmentation, the encoder expands the receptive field module is illustrated in image segmentation, the neoder 2) Multiscale Edge Restoration Module<br>
In the encoder-decoder architecture used for green tomato (ASPP) module is employ<br>
image segmentation, the encoder expands the receptive field<br>
incolume is illustrated in Fi<br>
through In the encoder-decoder architecture used for green tomato (ASPP) module is employed<br>
image segmentation, the encoder expands the receptive field module is illustrated in Fig.<br>
through downsampling, while the decoder resto image segmentation, the encoder expands the receptive field<br>
incough downsampling, while the decoder restores the<br>
incough incourse can result in the loss of semantic features at the handling the detail<br>
edges of green to through downsampling, while the decoder restores the<br>
image to its original size via upsampling. However, this<br>
different scales, enhancing<br>
process can result in the loss of semantic features at the<br>
dedges of green toma image to its original size via upsampling. However, this<br>
different scales, enhancing t<br>
process can result in the loss of semantic features at the<br>
indimig the details of green to<br>
edges of green tomatoes and the neglect process can result in the loss of semantic features at the<br>
information, which adversely affects segmentation accuracy.<br>
ASPP is a spatial attention mechanism for image<br>
ASPP is a spatial attention mechanism for image<br>
AS edges of green tomatoes and the neglect of contextual<br>information, which adversely affects segmentation accuracy.<br>ASPP is a spatial attention mechanism for image<br>segmentation accuracy.<br>ASPP is a spatial attention mechanism information, which adversely affects segmentation accuracy.<br>
ASPP is a spatial attention mechanism for image  $\frac{3}{2}$  *Loss Function*<br>
segmentation modeling designed to improve the capture of The goal of green tomato<br>
co ASPP is a spatial attention mechanism for image 3) Loss Function segmentation modeling designed to improve the capture of The goal of green contextual information. This module effectively captures the accurately recognize segmentation modeling designed to improve the capture of<br>contextual information. This module effectively captures the<br>corvolutions at different scales, thereby integrating these<br>corrollation, ensuring<br>convolutions at diffe contextual information. This module effectively captures the<br>features of green tomato images through dilated classification, ensuring a<br>convolutions at different scales, thereby integrating these<br>features to strengthen the features of green tomato images through dilated classification, ensuring<br>convolutions at different scales, thereby integrating these<br>the beat promotos and<br>semantic content of the images. The ASPP module consists imbalance convolutions at different scales, thereby integrating these<br>
features to strengthen the model's understanding of the the green tomatoes and<br>
semantic content of the images. The ASPP module consists<br>
imbalance in the number features to strengthen the model's understanding of the the green tomatoes and the sexuatoric content of the images. The ASPP module consists imbalance in the numbest content of the images. One 1x1 convolution primarily fo semantic content of the images. The ASPP module consists<br>
of five parallel branches: one 1x1 convolution primarily for<br>
extracting local information and reducing the number of<br>
extracting local information and reducing the of five parallel branches: one 1x1 convolution primarily for<br>extracting local information and reducing the number of<br>parameters; three 3x3 dilated convolutions with different<br>leading to an increasing dilation rates (6, 12, extracting local information and reducing the number of adequately learn the featur parameters; three 3x3 dilated convolutions with different leading to an increase dilation rates (6, 12, 18), allowing the convolution kern parameters; three 3x3 dilated convolutions with different<br>
dilation rates (6, 12, 18), allowing the convolution kernels to<br>
consequently, this issue seric<br>
cover a broader input area without increasing the number of<br>
the s dilation rates (6, 12, 18), allowing the convolution kernels to<br>consequently, this isse<br>cover a broader input area without increasing the number of<br>information. ASPP enlarges the receptive field through<br>convolutions with v cover a broader input area without increasing the number of<br>parameters, thereby aiding in capturing wider contextual<br>information. ASPP enlarges the receptive field through<br>convolutions with varying dilation rates, better c

To address this issue, the Atrous Spatial Pyramid Pooling<br>(ASPP) module is employed. The structure of the ASPP<br>module is illustrated in Fig. 5. This module effectively<br>captures detailed features through atrous convolution<br> To address this issue, the Atrous Spatial Pyramid Pooling<br>(ASPP) module is employed. The structure of the ASPP<br>module is illustrated in Fig. 5. This module effectively<br>captures detailed features through atrous convolution To address this issue, the Atrous Spatial Pyramid Pooling<br>(ASPP) module is employed. The structure of the ASPP<br>module is illustrated in Fig. 5. This module effectively<br>captures detailed features through atrous convolution<br> preserving the Atrons Spatial Pyramid Pooling<br>
(ASPP) module is employed. The structure of the ASPP<br>
module is illustrated in Fig. 5. This module effectively<br>
captures detailed features through atrous convolution<br>
differen J<br>
To address this issue, the Atrous Spatial Pyramid I<br>
(ASPP) module is employed. The structure of the<br>
module is illustrated in Fig. 5. This module effe<br>
captures detailed features through atrous conve<br>
different scales, *3* A address this issue, the Atrous Spatial Pyram<br> *SPP*) module is employed. The structure of<br> *3* Devices *a* illustrated in Fig. 5. This module<br> *8* ptures detailed features through atrous cofferent scales, enhancing t ondent and is issue, the Atrous Spatial Pyramid Pooling<br>SPP) module is employed. The structure of the ASPP<br>podule is illustrated in Fig. 5. This module effectively<br>ptures detailed features through atrous convolution<br>fferen To address this issue, the Atrous Spatial Pyramid Pooling<br>(ASPP) module is employed. The structure of the ASPP<br>module is illustrated in Fig. 5. This module effectively<br>captures detailed features through atrous convolution<br> To address this issue, the Atrous Spatial Pyramid Pooling<br>(ASPP) module is employed. The structure of the ASPP<br>module is illustrated in Fig. 5. This module effectively<br>captures detailed features through atrous convolution<br> To address this issue, the Atrous Spatial Pyramid Pooling<br>(ASPP) module is employed. The structure of the ASPP<br>module is illustrated in Fig. 5. This module effectively<br>captures detailed features through atrous convolution<br>

(ASPP) module is employed. The structure of the ASPP<br>module is illustrated in Fig. 5. This module effectively<br>captures detailed features through atrous convolution<br>different scales, enhancing the model's performance in<br>ha module is illustrated in Fig. 5. This module effectively<br>captures detailed features through atrous convolution<br>different scales, enhancing the model's performance in<br>handling the details of green tomato edges, and signific captures detailed features through atrous convolution<br>different scales, enhancing the model's performance in<br>handling the details of green tomato edges, and significantly<br>preserving edge details and significantly improving different scales, enhancing the model's performance in<br>handling the details of green tomato edges, and significantly<br>preserving edge details and significantly improving<br>segmentation accuracy.<br>3) Loss Function<br>The goal of g handling the details of green tomato edges, and significantly<br>preserving edge details and significantly improving<br>segmentation accuracy.<br>3) Loss Function<br>The goal of green tomato image segmentation is to<br>accurately recogni preserving edge details and significantly improving<br>segmentation accuracy.<br>3) Loss Function<br>The goal of green tomato image segmentation is to<br>accurately recognize green tomatoes through pixel-level<br>classification, ensuring segmentation accuracy.<br>3) Loss Function<br>The goal of green tomato image segmentation is to<br>accurately recognize green tomatoes through pixel-level<br>classification, ensuring a clear distinction from the<br>background. However, a 3) Loss Function<br>The goal of green tomato image segmentation is to<br>curately recognize green tomatoes through pixel-level<br>assification, ensuring a clear distinction from the<br>ekground. However, a significant size difference The goal of green tomato image segmentation is to<br>accurately recognize green tomatoes through pixel-level<br>classification, ensuring a clear distinction from the<br>background. However, a significant size difference between<br>the accurately recognize green tomatoes through pixel-level<br>classification, ensuring a clear distinction from the<br>background. However, a significant size difference between<br>the green tomatoes and the background, resulting in a classification, ensuring a clear distinction from the background. However, a significant size difference between the green tomatoes and the background, resulting in an imbalance in the number of pixels between the two cate

background. However, a significant size difference between<br>the green tomatoes and the background, resulting in an<br>imbalance in the number of pixels between the two<br>categories. This imbalance makes it difficult for the mode the green tomatoes and the background, resulting in an<br>imbalance in the number of pixels between the two<br>categories. This imbalance makes it difficult for the model to<br>adequately learn the features of the green tomatoes, o imbalance in the number of pixels between the two<br>categories. This imbalance makes it difficult for the model to<br>adequately learn the features of the green tomatoes, often<br>leading to an increase in false-negative predictio categories. This imbalance makes it difficult for the model to<br>adequately learn the features of the green tomatoes, often<br>leading to an increase in false-negative predictions.<br>Consequently, this issue seriously impacts the adequately learn the features of the green tomatoes, often leading to an increase in false-negative predictions.<br>Consequently, this issue seriously impacts the accuracy of the semantic segmentation of green tomatoes [29]. leading to an increase in false-negative predictions.<br>Consequently, this issue seriously impacts the accuracy of<br>the semantic segmentation of green tomatoes [29].<br>The Cross Entropy Loss (CE Loss) function is employed<br>to ca Consequently, this issue seriously impacts the accuracy of the semantic segmentation of green tomatoes [29]. The Cross Entropy Loss (CE Loss) function is employed to calculate the prediction accuracy for each pixel and the follows:



	<b>TABLE II</b> COMPARATIVE RESULTS OF THE IMPACT OF ASPP AND AG MODULES ON SWIN-UNET							
	<b>Base Model</b>	ASPP	AG	$PA(\%)$	$Dice(\% )$	$IoU(\%)$		
		$\times$	$\times$	93.0	86.5	76.3		
	Swin-Unet	$\times$		96.0	87.6	77.9		
			$\times$	96.8	90.2	82.2		
	Swin-Unet			97.5	92.4	85.9		
$Loss_{ce} = -\frac{1}{N} \sum_{i=1}^{N} [y_i log(y_i') + (1 - y_i) log(1 - y_i')]$			(3)		balance between Cross Entropy Loss and Dice Loss.	<b>III. RESULTS AND ANALYSIS</b>		
To effectively address the imbalance problem, Dice Loss is adopted. This loss function calculates the loss by comparing the similarity between the predicted probabilities and the true labels, making it particularly suitable for addressing category imbalance. Dice Loss ensures that the model optimizes the prediction of frequent categories while also noving ettention to infragment estageries. Consequently				To better validate the effectiveness of the model for green tomato segmentation, a series of experiments were conducted in this study. The experimental details were meticulously described, and the results were compared and analyzed. During the training process, the optimal model was selected and applied to the validation set to facilitate a				

$$
Loss_{ce} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(y_i') + (1 - y_i) \log(1 - y_i')] \qquad (3)
$$
 balance b

Swin-Unet<br>
<sup>x</sup> 96.0 87.6<br>
<br>
Swin-Unet<br>
<sup>7</sup> 97.5 92.4<br>
<br>
Loss<sub>ce</sub> =  $-\frac{1}{N}\sum_{i=1}^{N} [y_i \log(y'_i)+(1-y_i) \log(1-y'_i)]$  (3)<br>
To effectively address the imbalance problem, Dice Loss<br>
is adopted. This loss function calculates the loss b Swin-Unet<br>  $\angle$  50.0 8/.6 96.0 8/.6<br>
Swin-Unet<br>  $\angle$  7 × 96.8 90.2<br>
Swin-Unet<br>  $\angle$  7 × 96.8 90.2<br>
Swin-Unet<br>  $\angle$  7 × 96.8 90.2<br>
Sum-Unet<br>  $\angle$  7 × 96.8 90.2<br>
Sum-Unet<br>  $\angle$  7 × 96.8 90.2<br>
Sum-Unet<br>  $\angle$  97.5 92.4<br>
Sum-U Swin-Unet v × 96.8 90.2<br>
Swin-Unet v 97.5 92.4<br>  $Loss_{ce} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(y'_i)+(1-y_i) \log(1-y'_i)]$  (3)<br>
To effectively address the imbalance problem, Dice Loss<br>
is adopted. This loss function calculates the loss by<br>
comparing Swin-Unet **V** 97.5<br>  $Loss_{ce} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(y_i') + (1-y_i) \log(1-y_i')]$  (3)<br>
To effectively address the imbalance problem, Dice Loss<br>
is adopted. This loss function calculates the loss by<br>
comparing the similarity between the Swin-Unet<br>  $Loss_{ce} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(y'_i) + (1-y_i) \log(1-y'_i)]$  (3)<br>
To effectively address the imbalance problem, Dice Loss<br>
is adopted. This loss function calculates the loss by<br>
to better validate the effect<br>
comparing the Loss<sub>ce</sub> =  $-\frac{1}{N}\sum_{i=1}^{N} [y_i \log(y'_i)+(1-y_i)\log(1-y'_i)]$  (3)<br>
To effectively address the imbalance problem, Dice Loss<br>
is adopted. This loss function calculates the loss by<br>
comparing the similarity between the predicted probabi To encertrely attention and protectively attention content the protection calculates the loss by To better validate the effection aparamig the similarity between the predicted probabilities conducted in this study. The di is adopted. This loss function calculates the loss by<br>comparing the similarity between the predicted probabilities<br>and the true labels, making it particularly suitable for<br>addressing category imbalance. Dice Loss ensures

*L* oss 
$$
_{\text{dice}} = 1 - \frac{2 \sum_{i=1}^{N} y_i^T y_i}{\sum_{i=1}^{N} y_i^T + \sum_{i=1}^{N} y_i}
$$
 (4) proposed model in this s  
\n*A. Experimental Environ* The experimental set

$$
\frac{\partial Loss_{dice}}{\partial y'_i} = -\frac{2 y_i^2}{(y'_i + y_i)^2}
$$
 (5) NVIDIA

model optimizes the prediction of frequent categories while<br>
use selected and applied<br>
it effectively mitigates the model's bias toward the<br>
decorative experiment<br>
background in green tomato segmentation.<br>
Loss  $\lim_{\delta x \to$ also paying attention to infrequent categories. Consequently,<br>
it effectively mitigates the model's bias toward the<br>
background in green tomato segmentation.<br>
Loss  $\lim_{dx} = 1 - \frac{2 \sum_{i=1}^{x} y_i^2 y_i}{\sum_{i=1}^{x} y_i^2 + \sum_{i=1}^{y} y$ t effectively mitigates the model's bias toward the<br>
anackground in green tomato segmentation.<br>
Loss  $\frac{1}{\omega_{\text{new}}} = 1 - \frac{2 \sum_{i=1}^{N} y_i^* y_i}{\sum_{i=1}^{N} y_i^* + \sum_{i=1}^{N} y_i}$  (4) proposed model in this study<br>
The gradient form background in green tomato segmentation.<br>
Loss  $\omega_c = 1 - \frac{2 \sum_{i=1}^{8} y_i^2 y_i}{\sum_{i=1}^{8} y_i^2 + \sum_{i=1}^{8} y_i}$  (4) proposed model in this stu-<br>
The gradient form of the Dice Loss is complex, and its<br>
The experimental Environ T *L* oss  $\frac{z}{4\alpha} = 1 - \frac{2 \sum_{i=1}^{x} y_i' y_i}{\sum_{i=1}^{x} y_i' + \sum_{i=1}^{x} y_i}$  (4) proposed model in t<br>
The gradient form of the Dice Loss is complex, and its<br>
formula is as follows:<br>  $\frac{\partial Loss_{\text{dec}}}{\partial y_i'} = - \frac{2 y_i^2}{(y_i' + y_i)^2}$  (5) *L* oss  $\frac{d}{dx} = 1 - \frac{2}{\sum_{i=1}^{n} y_i} + \sum_{i=1}^{n} y_i$ <br>
The gradient form of the Dice Loss is complex, and its<br>
formula is as follows:<br>  $\frac{\partial Loss_{dice}}{\partial y'_i} = -\frac{2y_i^2}{(y'_i + y_i)^2}$ <br>
(5) NVIDIA A30, with C<br>
Based on (5), it can be **Example 1**  $\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} y_i$ <br>
The gradient form of the Dice Loss is complex, and its<br>
formula is as follows:<br>
formula is as follows:<br>  $\frac{\partial Loss_{dice}}{\partial y'_i} = -\frac{2y_i^2}{(y'_i + y_i)^2}$ <br>
(5) NVIDIA A30, with CUDA<br>
run using The gradient form of the Dice Loss is complex, and its<br>formula is as follows:<br> $\frac{\partial Loss_{dice}}{\partial y'_i} = -\frac{2 y_i^2}{(y'_i + y_i)^2}$  (5) NVI<br>Based on (5), it can be inferred that, in extreme scenario B. 1<br>when the values of andare very s  $\frac{\partial Loss_{\text{dec}}}{\partial y_i^r} = -\frac{2y_i^2}{(y_i^r + y_i)^2}$  (5) NVIDIA A30, with CUDA v<br>
Based on (5), it can be inferred that, in extreme scenario B. Parameter Settings<br>
when the values of andare very small, the gradient values of the i  $\frac{\partial y_i}{\partial y_i} = -\frac{y_i}{(y_i + y_i)^2}$  run using Python version<br>Based on (5), it can be inferred that, in extreme scenario B. Parameter Settings<br>when the values of andare very small, the gradient values Prior to inputting int<br>ma Based on (5), it can be inferred that, in extreme scenario B. Parameter Settings<br>when the values of andare very small, the gradient values Their coefficient to may become very large, potentially leading to more<br>may become Bessou on (5), it can be interested and the values of andare very small, the gradient values of ensure with the values may become very large, potentially leading to more were uniformly fixed at unstable training. To comp

$$
Loss = \alpha Loss_{ce} + (1 - \alpha) Loss_{dice}
$$
 (6)

2  $\sum_{i=1}^{n} y_i' y_i$  (4) proposed model in this study.  $f(A(9) = B(16))$ <br>  $96.0 = 86.5$   $76.3$ <br>  $96.0 = 87.6$   $77.9$ <br>  $96.8 = 90.2$   $82.2$ <br>  $97.5 = 92.4$   $85.9$ <br>
balance between Cross Entropy Loss and Dice Loss.<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the 93.0 86.5 76.3<br>
96.0 87.6 77.9<br>
96.8 90.2 82.2<br>
97.5 92.4 85.9<br>
balance between Cross Entropy Loss and Dice Loss.<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the model for green<br>
tomato segmentatio 96.0 87.6 77.9<br>
96.8 90.2 82.2<br>
97.5 92.4 85.9<br>
balance between Cross Entropy Loss and Dice Loss.<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the model for green<br>
tomato segmentation, a series of e 96.8 90.2 82.2<br>
97.5 92.4 85.9<br>
balance between Cross Entropy Loss and Dice Loss.<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the model for green<br>
tomato segmentation, a series of experiments were<br> 96.8 90.2 82.2<br>
97.5 92.4 85.9<br>
balance between Cross Entropy Loss and Dice Loss.<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the model for green<br>
tomato segmentation, a series of experiments were<br> 97.5 92.4 85.9<br>
balance between Cross Entropy Loss and Dice Loss.<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the model for green<br>
tomato segmentation, a series of experiments were<br>
conducted in thi 97.5 92.4 85.9 <br>
balance between Cross Entropy Loss and Dice Loss.<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the model for green<br>
tomato segmentation, a series of experiments were<br>
conducted in t **Experimental configurations of the system of the performance of the model for green**<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the model for green<br>
tomato segmentation, a series of experimental d balance between Cross Entropy Loss and Dice Loss.<br>
III. RESULTS AND ANALYSIS<br>
To better validate the effectiveness of the model for green<br>
tomato segmentation, a series of experiments were<br>
conducted in this study. The exp III. RESULTS AND ANALYSIS<br>To better validate the effectiveness of the model for gr<br>tomato segmentation, a series of experiments w<br>conducted in this study. The experimental details w<br>meticulously described, and the results The effectiveness of the model for green<br>
To better validate the effectiveness of the model for green<br>
mato segmentation, a series of experiments were<br>
reticulously described, and the results were compared and<br>
alyzed. Dur To better validate the effectiveness of the model for green<br>tomato segmentation, a series of experiments were<br>conducted in this study. The experimental details were<br>meticulously described, and the results were compared and tomato segmentation, a series of experiments were<br>conducted in this study. The experimental details were<br>meticulously described, and the results were compared and<br>analyzed. During the training process, the optimal model<br>wa conducted in this study. The experimental details were<br>meticulously described, and the results were compared and<br>analyzed. During the training process, the optimal model<br>was selected and applied to the validation set to fa meticulously described, and the results were compared and<br>analyzed. During the training process, the optimal model<br>was selected and applied to the validation set to facilitate a<br>comparative evaluation of the experimental o **Example 12**<br>
Was selected and applied to the validation set to facil<br>
comparative evaluation of the experimental out<br>
Comparative experiments were conducted under ide<br>
experimental configurations to assess the performance

The experimental setup is based on the Ubuntu 18.04 run using Python version 3.7 and PyTorch version 1.12. mparative evaluation of the experimental outcomes.<br>
mparative experiments were conducted under identical<br>
perimental configurations to assess the performance of the<br>
poposed model in this study.<br>
Experimental Environment<br>

batch size of 2, window size of 7, and patch size of 4x4. The learning rate curve variation is shown in Figure 6. Using the Comparative experiments were conducted under identical<br>experimental configurations to assess the performance of the<br>proposed model in this study.<br>A. Experimental Environment<br>The experimental setup is based on the Ubuntu 18 experimental configurations to assess the performance of the<br>proposed model in this study.<br>A. Experimental Environment<br>The experimental setup is based on the Ubuntu 18.04<br>64-bit system, utilizing the deep learning framewor proposed model in this study.<br>
A. Experimental Environment<br>
The experimental setup is based on the Ubuntu 18.04<br>
64-bit system, utilizing the deep learning framework<br>
PyTorch. The GPU used for the experiments is a 24GB<br>
NV A. Experimental Environment<br>The experimental setup is based on the Ubuntu 18.04<br>64-bit system, utilizing the deep learning framework<br>PyTorch. The GPU used for the experiments is a 24GB<br>NVIDIA A30, with CUDA version 11.4. A The experimental entwolutions<br>
The experimental setup is based on the Ubuntu 18.04<br>
64-bit system, utilizing the deep learning framework<br>
64-bit system, utilizing the deep learning framework<br>
PyTorch. The GPU used for the The experimental setup is based on the countal 18.04<br>
64-bit system, utilizing the deep learning framework<br>
64-bit system, utilizing the deep learning framework<br>
PyTorch. The GPU used for the experiments is a 24GB<br>
NVIDIA beyond system, universign the decept canning manitowite<br>PyTorch. The GPU used for the experiments is a 24GB<br>NVIDIA A30, with CUDA version 11.4. All models were<br>run using Python version 3.7 and PyTorch version 1.12.<br>B. Par Level the GF of side for the experiments is a 2-volume of the experimental is a 2-volume variant physion 1.14. All models were run using Python version 3.7 and PyTorch version 1.12.<br>B. Parameter Settings<br>Prior to inputtin France training parameters, which curve training parameter Settings<br>
From using Python version 3.7 and PyTorch version 1.12.<br>
B. Parameter Settings<br>
Prior to inputting into the training network, image sizes<br>
were uniforml B. Parameter Settings<br>B. Parameter Settings<br>Prior to inputting into the training network, image sizes<br>were uniformly fixed at (640,640). Pre-trained weights from<br>the ImageNet dataset, specifically swin\_tiny\_patch4<br>\_window *B. Parameter Settings*<br>Prior to inputting into the training netword<br>were uniformly fixed at  $(640,640)$ . Pre-traine<br>the ImageNet dataset, specifically sweed with an initial<br>of 0.01, momentum of 0.9, weight decay of<br>stoch



e)+AG+ASPP

Fig.7. Comparative Visualization of Ablation Study Results.<br>
C. Evaluation Metrics<br>
To evaluate the segmentation accuracy of the optimized<br>
Swin-Unet algorithm for green tomatoes, metrics such as<br>
Precision, Recall, Pixel

$$
Precision = \frac{TP}{TP + FP} \times 100\%
$$
 (8) The Dice coeffi-  
between two sets,

$$
Recall = \frac{TP}{TP + FN} \times 100\%
$$
 (9)

SPP

\n
$$
Recall = \frac{TP}{TP + FN} \times 100\%
$$
 (9)

\nPA denotes the ratio of the number of correct predictions for all pixel classes to the total number of pixels.

\n
$$
PA = \frac{TP + TN}{TP + FP + TN + FN}
$$
 (10)

\nThe Dice coefficient is used to calculate the similarity between two sets, as shown in equation (11):

The Dice coefficient is used to calculate the similarity

# **Volume 32, Issue 11, November 2024, Pages 2114-2126**

$$
Dice = \frac{2TP}{FP + 2TP + FN}
$$
 (11) without adding extra pa  
reflected in the model's im

**Engineering Letters**<br>
Dice =  $\frac{2TP}{FP + 2TP + FN}$  (11) without adding extra param<br>
The Intersection over Union (IoU) represents the ratio of the edge details and overall semaintersection to the union of two sets, as illustra **Engineering Letters**<br>
Dice  $= \frac{2TP}{FP + 2TP + FN}$  (11) without adding extra parameters<br>
The Intersection over Union (IoU) represents the ratio of the dege details and overall semant<br>
intersection to the union of two sets, as Dice  $=$   $\frac{2TP}{FP + 2TP + FN}$ <br>The Intersection over Union (IoU) represents the<br>intersection to the union of two sets, as il<br>equation (12):<br> $IoU = \frac{TP}{FP + TP + FN}$ <br>This paper defines the metric using a confu-

$$
IoU = \frac{TP}{FP + TP + FN}
$$
 the base model, compa  
PA, Dice, and IoU nor

**Engineering Letters**<br>
Dice  $= \frac{2TP}{FP + 2TP + FN}$  (11) without adding extra proposed in the model's in<br>
intersection over Union (IoU) represents the ratio of the edge details and overall sintersection to the union of two sets **Engineering Letters**<br> *Dice* =  $\frac{2TP}{FP + 2TP + FN}$  (11) without adding extraction over Union (IoU) represents the ratio of the dge details and over intersection to the union of two sets, as illustrated in the mode relations Dice  $=$   $\frac{2TP}{FP + 2TP + FN}$ <br>
The Intersection over Union (IoU) represents the ratio of the edge details and overall semi-<br>
intersection to the union of two sets, as illustrated in the model's improducing both<br>
edge details *Dice* =  $\frac{2TP}{FP + 2TP + FN}$  (11) without adding extra para<br>
The Intersection over Union (IoU) represents the ratio of the edge details and overall sem<br>
intersection to the union of two sets, as illustrated in the model's im *Dice* =  $\frac{2TP}{FP + 2TP}$  (11) without adding extra par<br>
The Intersection over Union (IoU) represents the ratio of the<br>
intersection to the union of two sets, as illustrated in the model's implement<br>
equation (12):<br>  $IoU = \frac{$ *Dice* =  $\frac{2IP}{FP + 2TP + FN}$  (11) without adding extra in the model's intersection over Union (IoU) represents the ratio of the edge details and overall intersection to the union of two sets, as illustrated in more precise e samples. *D. Ablation Study*<br>*D. Ablation Study*<br>To verify the effectiveness of these two equation (12):<br>  $F_{10}$   $F_{11}$   $F_{12}$   $F_{13}$   $F_{14}$   $F_{15}$   $F_{16}$   $F_{17}$   $F_{18}$   $F_{19}$   $F_{18}$   $F_{19}$   $F_{18}$   $F_{19}$   $F_{19}$   $F_{18}$   $F_{19}$   $F_{19}$   $F_{19}$   $F_{19}$   $F_{19}$   $F_{19}$   $F_{19}$   $F_{19}$   $F_{19}$   $F_{19$  $IoU = \frac{TP}{FP + TP + FN}$  (12) the base model, comp<br>
This paper defines the metric using a confusion matrix, points, respectively.<br>
categorizing green tomato samples based on the relationship<br>
between predicted values and actual v  $I = \frac{I_0}{F + TP + FN}$  PA, Dice, and IoU increased b<br>This paper defines the metric using a confusion matrix,<br>categorizing green tomato samples based on the relationship<br>between predicted values and actual values into four<br>the This paper defines the metric using a confusion matrix,<br>
categorizing green tomato samples based on the relationship<br>
between predicted values and actual values into four<br>
tomato image segmentation results mo<br>
positive (F categorizing green tomato samples based on the relationship<br>between predicted values and actual values into four<br>detections of the setting of the compare<br>categories: true positive (TP): actual positive samples; false the

between predicted values and actual values into four<br>
between predicted values and actual positive samples; false  $\frac{1}{2}$  and order to<br>
positive (FP): false positive samples; true negative (TN):<br>
actual negative samples Examples: the positive (TP): actual positive samples; false<br>
actual negative (TP): false positive samples; the negative (TN):<br>
actual negative samples; false negative (TN): False negative ablation results meant<br>
actual ne

Fraction experiment<br>
actual negative samples; false negative (FN): False negative<br>
samples.<br>
D. Ablation Study<br>
To verify the effectiveness of these two structures, the<br>
poptimized Swin-UNet algorithm and the original Swi samples.<br>
Samples.<br>
Contriguent of the comparison of the subset of the subset of the original Swin<br>
To verify the effectiveness of these two structures, the<br>
process and comparability the criginal Swin-UNet<br>
algorithm wer entiative the Higher Contention accuracy for green tomators and more optimized Swin-UNet algorithm and the original Swin-UNet algorithm and the original Swin-UNet algorithm were evaluated in the Green Tomato dataset. To an *D. Ablation Study*<br>
To verify the effectiveness of these two structures, the<br>
algorithm and the original Swin-UNe<br>
algorithm and the original Swin-UNE<br>
algorithm were valuated in the Green Tomato dataset. To<br>
algorithms e To verify the effectiveness of these two structures, the<br>
optimized Swin-UNet algorithm and the original Swin-UNet<br>
algorithm were evaluated in the Green Tomato dataset. To<br>
ended and clearer, and the sum-<br>
ended in the G optimized Swin-UNet algorithm and the original Swin-UNet<br>algorithm were evaluated in the Green Tomato dataset. To<br>and clearer, and the segn<br>ensure fairness and comparability the experimental settings<br>much closer to that of algorithm were evaluated in the Green Tomato dataset. To and clearer, and the segmensure fairness and comparability the experimental settings much closer to that of the read algorithms, as detailed in Table II. the introdu ensure fairness and comparability the experimental settings<br>
much closer to that of t<br>
aros all algorithms, as detailed in Table II. which fully proves<br>
As shown in Table II, the introduction of the AG module<br>
into the bas and hyperparameter configurations were kept consistent<br>
are segmentation edges of<br>
across all algorithms, as detailed in Table II.<br>
As shown in Table II, the introduction of the AG module<br>
into the base model by 3.0, 1.1, across all algorithms, as detailed in Table II. which fully proves that the AS shown in Table II, the introduction of the AG module of the effectively improves green tom<br>
into the base model improves the PA, Dice, and IoU As shown in Table II, the introduction of the AG module<br>
into the base model improves the PA, Dice, and IoU of the<br>
model by 3.0, 1.1, and 1.6 percentage points, respectively,<br>
compared to the original model. This is gini

Example 19 Letters<br>without adding extra parameters. This enhancement is<br>reflected in the model's improved performance in capturing<br>edge details and overall semantic understanding, leading to<br>more precise edge segmentation Figures<br>is the model of the model's improved performance in capturing<br>reflected in the model's improved performance in capturing<br>edge details and overall semantic understanding, leading to<br>more precise edge segmentation of Example 1. **Letters**<br>
without adding extra parameters. This enhancement is<br>
reflected in the model's improved performance in capturing<br>
edge details and overall semantic understanding, leading to<br>
more precise edge segmen **Example 18 Example 15 Example 15 Example 10**<br>The preflected in the model's improved performance in capturing<br>edge details and overall semantic understanding, leading to<br>more precise edge segmentation of green tomatoes.<br>Fi

points, respectively. This further validates that both Finally, by introducing both ASPP and AG modules into **Example 18**<br>
is the base model of the model's improved performance in capturing<br>
edge details and overall semantic understanding, leading to<br>
more precise edge segmentation of green tomatoes.<br>
Finally, by introducing both Figure 1. **Letters**<br>
without adding extra parameters. This enhancement is<br>
reflected in the model's improved performance in capturing<br>
edge details and overall semantic understanding, leading to<br>
more precise edge segmenta **Figure 12**<br> **Exercise 12**<br> **Exerce** in the model's improved performance in capturing<br>
edge details and overall semantic understanding, leading to<br>
more precise edge segmentation of green tomatoes.<br>
Finally, by introducing **Exert Exercise 15**<br> **Exert Exercise 15**<br> **Exerce** model and overall simproved performance in capturing<br>
edge details and overall semantic understanding, leading to<br>
more precise edge segmentation of green tomatoes.<br>
Final without adding extra parameters. This enhancement<br>reflected in the model's improved performance in ca<br>edge details and overall semantic understanding, lea<br>more precise edge segmentation of green tomatoes.<br>Finally, by intr In thout adding extra parameters. This enhancement is<br>
elected in the model's improved performance in capturing<br>
ge details and overall semantic understanding, leading to<br>
ore precise edge segmentation of green tomatoes.<br> without adding extra parameters. This enhancement is<br>reflected in the model's improved performance in capturing<br>edge details and overall semantic understanding, leading to<br>more precise edge segmentation of green tomatoes.<br>

**Examples:**<br> **Examples:** The present to the base model in the base model in the base model in the present of the present of the base model in the present of the present of the properties and the original Swin-UNet algorit without adding extra parameters. This enhancement is<br>reflected in the model's improved performance in capturing<br>edge details and overall semantic understanding, leading to<br>more precise edge segmentation of green tomatoes.<br> reflected in the model's improved performance in capturing<br>edge details and overall semantic understanding, leading to<br>more precise edge segmentation of green tomatoes.<br>Finally, by introducing both ASPP and AG modules into edge details and overall semantic understanding, leading to<br>more precise edge segmentation of green tomatoes.<br>Finally, by introducing both ASPP and AG modules into<br>the base model, compared to the original model, the model' more precise edge segmentation of green tomatoes.<br>
Finally, by introducing both ASPP and AG modules into<br>
the base model, compared to the original model, the model's<br>
PA, Dice, and IoU increased by 4.5, 5.9, and 9.6 percen Finally, by introducing both ASPP and AG modules into<br>the base model, compared to the original model, the model's<br>PA, Dice, and IoU increased by 4.5, 5.9, and 9.6 percentage<br>points, respectively. This further validates tha the base model, compared to the original model, the model's<br>
PA, Dice, and IoU increased by 4.5, 5.9, and 9.6 percentage<br>
points, respectively. This further validates that both<br>
proposed modules effectively enhance accurac PA, Dice, and IoU increased by 4.5, 5.9, and 9.6 percentage<br>points, respectively. This further validates that both<br>proposed modules effectively enhance accuracy in green<br>tomato image segmentation.<br>In order to compare the e points, respectively. This further validates that both proposed modules effectively enhance accuracy in green tomato image segmentation.<br>In order to compare the effect of each module on the segmentation results more intuit proposed modules effectively enhance accuracy in green<br>tomato image segmentation.<br>In order to compare the effect of each module on the<br>segmentation results more intuitively, the results of the<br>ablation experiments are visu tomato image segmentation.<br>
In order to compare the effect of each module on the<br>
segmentation results more intuitively, the results of the<br>
ablation experiments are visualized in this paper, and the<br>
visualization results In order to compare the effect of each module on the segmentation results more intuitively, the results of the ablation experiments are visualized in this paper, and the visualization results are shown in Fig. 7. As can be **E.** Segmentation results and the model in this paper, and in this paper and introduction results are shown in Fig. 7. As can be seem Fig. 7, the original Swin-Unet model exhibits issue as unclear segmentation edges and ta Example 12 and the serve the performance of the algorithms using the Green Species as unclear segmentation edges and target leakage, etc.<br>
Specifical swin-Unet model exhibits issues the as unclear segmentation edges and ta from Fig 7, the original Swin-Unet model exhibits issues<br>such as unclear segmentation edges and target leakage, etc.<br>However, with the gradual introduction of modules such as<br>ASPP and AG, the target contour becomes more a such as unclear segmentation edges and target leakage, etc.<br>However, with the gradual introduction of modules such as<br>ASPP and AG, the target contour becomes more accurate<br>and clearer, and the segmentation effect of the mo

However, with the gradual introduction of modules such as<br>ASPP and AG, the target contour becomes more accurate<br>and clearer, and the segmentation effect of the model is<br>much closer to that of the real labels, and the pheno ASPP and AG, the target contour becomes more accurate<br>and clearer, and the segmentation effect of the model is<br>much closer to that of the real labels, and the phenomenon of<br>unclear segmentation edges of the target leakage and clearer, and the segmentation effect of the model is<br>much closer to that of the real labels, and the phenomenon of<br>unclear segmentation edges of the target leakage is reduced,<br>which fully proves that the model proposed much closer to that of the real labels, and the phenomenon of<br>unclear segmentation edges of the target leakage is reduced,<br>which fully proves that the model proposed model<br>effectively improves green tomato image segmentati unclear segmentation edges of the target leakage is reduced,<br>which fully proves that the model proposed model<br>effectively improves that the model proposed model<br>effectively improves green tomato image segmentation.<br>E. Segm which fully proves that the model proposed model<br>effectively improves that the model proposed model<br>effectively improves green tomato image segmentation.<br>*E. Segmentation Results*<br>To further analyze the performance of the criteria. optimized Swin-Unet is compared contemporary and advanced sem<br>
algorithms using the Green Tor<br>
algorithms include<br>
DeepLabv3 [32], PSPNet [33], DAN<br>
E ISA-Net [36], DPT [37], OCRNet [38]<br>
e experiments were conducted under



# **Engineering Letters Engineering Eccuers**

<b>Engineering Letters</b>													
	<b>TABLE IV</b> COMPARISON OF THE NUMBER OF PARAMETERS AND FLOPS COMPUTATIONAL COMPLEXITY OF MODELS. INPUT SIZE: (640,640).												
Method	Deeplabv3+	Deeplabv3	Pspnet	Danet	Knet	Isanet	Dpt	Ocrnet	Beit	Ours			
Params/M	41.216	65.74	46.602	47.485	60.412	35.344	110	12.067	72.137	27.55			
GFLOPs/G	276	422	279	338	320	235	360	82.902	437	116.15			
	I Overlapping <b>II</b> Block the tomato tomato image image		III Overhead shot of tomato image (a) Original images of Tomato		IV Distant view tomato image		V Backlighting tomato image		VI Nighttime tomato image				













VBacklighting VI Nighttime tomato image









































# **Engineering Letters Engineering Eccuers**



Fig.8. Comparative Visualization of Experimental Result<br>
(1) Ours<br>
Fig.8. Comparative Visualization of Experimental Result<br>
(1) Ours<br>
Evaluation metrics compared to the comparison experiments were balance betwe<br>
conducted All comparation of Experimental Result<br>All comparison algorithms were trained and tested on the demonstrates strong overall p<br>the pen tomato dataset, and the comparison experiments were balance between model can<br>ducted usi Fig.8. Comparative Visualization of Experimental Result<br>(1) Ours<br>
Fig.8. Comparative Visualization of Experimental Result<br>
All comparison algorithms were trained and tested on the<br>
demonstrates strong or<br>
conducted using Fig.8. Comparative Visualization of Experimental Result<br>
All comparison algorithms were trained and tested on the<br>
demonstrates strong overall per<br>
green tomato dataset, and the comparison experiments were balance between

Fig.8. Comparative Visualization of Experimental Result<br>
II courses<br>
and the comparison agent and tested on the demonstrates strong overall g<br>
green tomato dataset, and the comparison experiments were balance between mode All comparison algorithms were trained and tested on the demonstrates strong overall p<br>green tomato dataset, and the comparison experiments were balance between model ca<br>conducted using MMsegmentation version 1.2.2. The e All comparison algorithms were trained and tested on the demonstrates strong overa<br>green tomato dataset, and the comparison experiments were balance between model<br>conducted using MMsegmentation version 1.2.2. The efficien green tomato dataset, and the comparison experiments were balance between model ca<br>conducted using MMsegmentation version 1.2.2. The efficiency.<br>segmentation results of each model are presented in Table III. From this anal conducted using MMsegmentation version 1.2.2. The efficiency.<br>
segmentation results of each model are presented in Table III. From this analysis, it is<br>
It is observed that the optimized Swin-Unet algorithm Swin-Unet algor segmentation results of each model are presented in Table III. From this analysis, it it is observed that the optimized Swin-Unet algorithm Swin-Unet algorithm demonstrates strong competitiveness across various across all

efficiency.

demonstrates strong overall performance, maintaining a balance between model capacity and computational efficiency.<br>From this analysis, it is evident that the optimized Swin-Unet algorithm demonstrates significant improvem demonstrates strong overall performance, maintaining a balance between model capacity and computational efficiency.<br>From this analysis, it is evident that the optimized Swin-Unct algorithm demonstrates significant improvem demonstrates strong overall performance, maintaining a balance between model capacity and computational efficiency.<br>From this analysis, it is evident that the optimized Swin-Unet algorithm demonstrates significant improvem demonstrates strong overall performance, maintaining a balance between model capacity and computational efficiency.<br>From this analysis, it is evident that the optimized Swin-Unet algorithm demonstrates significant improvem demonstrates strong overall performance, maintaining a balance between model capacity and computational efficiency.<br>From this analysis, it is evident that the optimized Swin-Unet algorithm demonstrates significant improvem demonstrates strong overall performance, maintaining a balance between model capacity and computational efficiency.<br>From this analysis, it is evident that the optimized Swin-Unet algorithm demonstrates significant improvem demonstrates strong overall performance, maintaining a<br>balance between model capacity and computational<br>efficiency.<br>From this analysis, it is evident that the optimized<br>Swin-Unet algorithm demonstrates significant improvem demonstrates strong overall performance, maintaining a balance between model capacity and computational efficiency.<br>From this analysis, it is evident that the optimized Swin-Unet algorithm demonstrates significant improvem balance between model capacity and computational efficiency.<br>
From this analysis, it is evident that the optimized<br>
Swin-Unet algorithm demonstrates significant improvements<br>
across all assessment metrics. Although issues efficiency.<br>From this analysis, it is evident that the optimized<br>Swin-Unet algorithm demonstrates significant improvements<br>across all assessment metrics. Although issues such as target<br>miss-detection and unclear segmentati From this analysis, it is evident that the optimized<br>Swin-Unet algorithm demonstrates significant improvements<br>across all assessment metrics. Although issues such as target<br>miss-detection and unclear segmentation edges are

**Engineering Letters**<br>While other algorithms face challenges such as target edge segmentation of gree<br>ssing and misclassification when dealing with fruit parts improving the success rate<br>der occlusion and overlapping scena **Engineering Letters**<br>While other algorithms face challenges such as target edge segmentation of gree<br>missing and misclassification when dealing with fruit parts improving the success rate<br>under occlusion and overlapping s Engineering Letters<br>
While other algorithms face challenges such as target<br>
missing and misclassification when dealing with fruit parts<br>
under occlusion and overlapping scenarios, the improved<br>
Swin-Unet algorithm effectiv **Engineering Letters**<br>
While other algorithms face challenges such as target edge segmentation of green<br>
missing and misclassification when dealing with fruit parts improving the success rate o<br>
under occlusion and overlap **Engineering Letters**<br>
While other algorithms face challenges such as target<br>
edge segmentation of green<br>
missing and misclassification when dealing with fruit parts<br>
improving the success rate c<br>
under occlusion and overl **Engineering Letters**<br>
While other algorithms face challenges such as target edge segmentation of green<br>
missing and misclassification when dealing with fruit parts improved Deverall, the optimized Swin-Un<br>
Swin-Unet algor **Engineering Letters**<br>
While other algorithms face challenges such as target edge segmentation of<br>
missing and misclassification when dealing with fruit parts improving the success in<br>
under occlusion and overlaping scenar **Engineering Letters**<br>
While other algorithms face challenges such as target edge segmentation of gree<br>
missing and misclassification when dealing with fruit parts improving the success rate<br>
under occlusion and overlappin **Engineering Letters**<br>
While other algorithms face challenges such as target edge segmentation of green<br>
missing and misclassification when dealing with fruit parts<br>
improving the success rate of<br>
under occlusion and overl **Engineering Letters**<br>
While other algorithms face challenges such as target edge segmentation of<br>
missing and misclassification when dealing with fruit parts improving the success<br>
under occlusion and overlapping scenario While other algorithms face challenges such as target edge segmentation of green<br>missing and misclassification when dealing with fruit parts improving the success rate of<br>under occlusion and overlapping scenarios, the impr While other algorithms face challenges such as target edge segmentation<br>missing and misclassification when dealing with fruit parts improving the succe<br>under occlusion and overlapping scenarios, the improved Overall, the o While other algorithms face challenges such as target edge segmentation of missing and misclassification when dealing with fruit parts improving the success<br>sum-Unret algorithm effectively addresses these issues and segmen missing and misclassification when dealing with fruit parts<br>
improving the success rate of<br>
under occlusion and overlapping scenarios, the improved Overall, the optimized Swin-Une<br>
Swin-Unet algorithm effectively addresses under occlusion and overlapping scenarios, the improved Overall, the optimized Swin-Swin-Unet algorithm effectively addresses these issues and segmentation performance a<br>surpasses them in delineating fruit edge details. Th Swin-Unet algorithm effectively addresses these issues and<br>surpasses them in delineating fruit edge details. The<br>offering a theoretical referent<br>visualization of comparative experiments shows that the<br>green fruits. However surpasses them in delineating fruit edge details. The offering a theoretical reference fivisualization of comparative experiments shows that the green fruits. However, despite the improved algorithm excels in achieving cle visualization of comparative experiments shows that the green fruits. However, despite timproved algorithm excels in achieving clear boundary by the optimized Swin-Unet a<br>segmentation and maintaining the integrity of targe improved algorithm excels in achieving clear boundary by the optimized Swin-Uned<br>gementation and maintaining the integrity of target fruits semantic segmentation all,<br>when capturing green tomatoes from various angles, such segmentation and maintaining the integrity of target fruits semantic seguhen capturing green tomatoes from various angles, such as complexity r aerial and long-distance views. Compared to other algorithms, particularly p i en capturing green tomatoes from various angles, such as<br>
irial and long-distance views. Compared to other algorithms, particularly prominent in<br>
more accurately resolves these issues. Additionally, in processing or limite aerial and long-distance views. Compared to other algorithms, particularly prominent in<br>t more accurately resolves these issues. Additionally, in processing or limited<br>backlighting scenarios, the improved algorithm demonst it more accurately resolves these issues. Additionally, in processing or limited consequently the distributions over other models. The optimized Swin-Unet networks and exergential and swin-Unet algorithm provides remarkabl backlighting scenarios, the improved algorithm demonstrates<br>
secarch should focus on furthe<br>
significant advantages over other models. The optimized Swin-Unet networks and explor<br>
Swin-Unet algorithm provides remarkable im

significant advantages over other models. The optimized Swin-Unet networks and exportantions by designs. Such efforts aim to delivering more precise edge segmentation, reducing errors, on computational resource and greatl Swin-Unet algorithm provides remarkable improvements by<br>
designs. Such efforts aim to m<br>
delivering more precise edge segmentation, reducing errors, on computational resource<br>
experimental results validate the effectivene delivering more precise edge segmentation, reducing errors, on computational resources<br>
and greatly enhancing overall performance. These segmentation accuracy, thereb<br>
experimental results validate the effectiveness of Swi and greatly enhancing overall performance. These segmentation accuracy, the experimental results validate the effectiveness of Swin-Unet model to bette incorporating ASPP and AG structures to optimize the requirements, esp experimental results validate the effectiveness of Swin-Unet model to bet<br>
incorporating ASPP and AG structures to optimize the<br>
requirements, especially<br>
Swin-Unet algorithm specifically for green tomato performance and r incorporating ASPP and AG structures to optimize the requirements, especially<br>
Swin-Unet algorithm specifically for green tomato performance and resource<br>
Segmentation in controlled agricultural environments.<br>
This study d Swin-Unet algorithm specifically for green tomato performance and resource<br>segmentation in controlled agricultural environments.<br>This study demonstrates the refinement of the optimized<br>images win-Unet algorithm, which add segmentation in controlled agricultural environments.<br>
This study demonstrates the refinement of the optimized<br>
Swin-Unet algorithm, which addresses complex challenges include<br>
in facility agriculture applications. These c This study demonstrates the refinement of the optimized<br>
Swin-Unet algorithm, which addresses complex challenges<br>
in facility agriculture applications. These challenges include<br>
transitions between day and inglit environm Swin-Unet algorithm, which addresses complex challenges<br>
in facility agriculture applications. These challenges include<br>
variable lighting conditions, diverse shooting angles, and<br>
transitions between day and night enviro in facility agriculture applications. These challenges include<br>variable lighting conditions, diverse shooting angles, and<br>transitions between day and night environments. Green<br>tomato images were capabre capabre capabre ca variable lighting conditions, diverse shooting angles, and<br>
transitions between day and night environments. Green<br>
tomato images were captured under various conditions, such<br>
as occlusions, overlaps, and different lightin transitions between day and night environments. Green<br>
tomato images were captured under various conditions, such<br>
seview. Critical Reviews in<br>
eview. Critical Reviews in<br>
eview. Critical Reviews in<br>
eview. The step excel tomato images were captured under various conditions, such<br>
segmentation, overlaps, and different lighting scenarios, to<br>
evaluate the algorithm's adaptability to these practical issues.<br>
Seience, 2023, 14: 1121209.<br>
Seie as occlusions, overlaps, and different lighting scenarios, to [2] Tiwari J K, Singh A K, Behera<br>
evaluate the algorithm's adaptability to these practical issues. Some the experimental results indicate that the algorithm a evaluate the algorithm's daptability to these practical issues.<br>
The experimental results indicate that the algorithm achieves<br>
Econe, 2023, 14: 1121209.<br>
excellent segmentation accuracy across different scenarios,<br>
excerc The experimental results indicate that the algorithm achieves<br>
excellent segmentation accuracy across different scenarios,<br>
effectively managing images with varying lighting conditions<br>
and backgrounds while accurately id agriculture. while accurately identifying and <sup>[4]</sup><br>d or overlapping fruits. These tests<br>rm the algorithm's robustness,<br>iility, and theoretical sophistication,<br>fectiveness in real-world applications. [6]<br>sive performance in green tomat menting occluded or overlapping truits. These tests<br>
lengrithm's robustness, [5] Xiong Y, Ge Y, Grimstad L,<br>
lengrithm's robustness, harvesting robust design, d<br>
menstating its effectiveness in real-world applications,<br>
wo collectively confirm the algorithm's robustness, havesting robot: design, deen<br>enconstrained signal-word applications, evaluation. Jounal of Field Robot<br>Beyond its impressive performance in green tomato the expendent of th generalization capability, and theoretical sophistication,<br>
demonstrating its effectiveness in real-world applications. [6] Jia W, Zhang Y, Lian J, et al. App<br>
Beyond its impressive performance in green tomato<br>
systems, 20

demonstrating its effectiveness in real-world applications.<br>
Beyond its impressive performance in green tomato systems, 2020, 17(3)<br>
segmentation, the optimized Swin-Unet algorithm shows [7] Saranya N, Srinivass<br>
potentia Beyond its impressive performance in green tomato<br>
segmentation, the optimized Swin-Unet algorithm shows<br>
segmentation for broader applications, including fruit<br>
segmentation in similar environments, thus upporting the<br>
s segmentation, the optimized Swin-Unet algorithm shows [1] saing raditional machine learning the potential for broader applications, including fruit computational vision and Bio-Insequentation in similar environments, thus potential for broader applications, including fruit  $\frac{1}{2}$  computational Vision and Exegenentation in similar environments, thus supporting the  $\frac{1}{2}$  Computational Vision and Exegenentation of automation techniques Examentation in similar environments, thus supporting the<br>
wider adoption of automation techniques in facility-based [8] Fujinga T, Nakanishi T<br>
agriculture.<br>
IV. CONCLUSION [9] Marizuan Mat Dault by<br>
Exercise of green fr wider adoption of automation techniques in facility-based [8] Fujing DeplabV3+ for<br>agriculture.<br>
The agricultures are at the Dallenges posed by the similar color<br>
Features of green fruits and the background of branches an agriculture.<br>
IV. CONCLUSION<br>
Features of green fruits and the background of branches and<br>
features of green fruits and the ba For CONCLUSION<br>
Marizuana Mat Daud, Zulai<br>
Matersing the challenges posed by the similar color<br>
ready-to-harvest prediction of original<br>
features of green fruits and the background of branches and<br>
leaves, as well as issu For CONCLUSION<br>
Addressing the challenges posed by the similar color<br>
features of green fruits and the background of branches and<br>
leaves, as well as issues of occulusom, overlapping, and<br>
leaves, as well as issues of occ Addressing the challenges posed by the similar color<br>
features of green fruits and the background of branches and<br>
leaves, as well as issues of occlusion, overlapping, and<br>
leaves, as well as issues of occlusion, overlapp features of green fruits and the background of branches and<br>
leavers, as well as issues of occlusion, overlapping, and<br>
varying lighting conditions that complicate segmentation,<br>
where the counting methods for yield may<br> leaves, as well as issues of occlusion, overlapping, and [10] Häni N, Roy P, Isler V. A contring lighting conditions that complicate segmentation,<br>
this thesis utilizes a custom dataset of green tomatoes. The<br>
semantic se varying lighting conditions that complicate segmentation,<br>
this thesis utilizes a custom dataset of green tomatoes. The<br>
interational Robotics, 2020, 37(2): 26;<br>
interational Configure alternational Configure alternationa this thesis utilizes a custom dataset of green tomatoes. The [11] Bargoti S, Underwood J. Desemantic segmentation algorithm Swin-Unet is optimized by<br>
integrating the Atrous Spatial Pyramid Pooling (ASPP)  $2017:3626-3633$ 

g Letters<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation performance and strong g Letters<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation performance and strong **g Letters**<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation performance and stro g Letters<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation performance and strong **g Letters**<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation performance and stro g Letters<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation performance and strong **g Letters**<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation performance and stro **g Letters**<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation performance and stro **g Letters**<br>edge segmentation of green fruits, thereby effectively<br>improving the success rate of picking in real scenarios.<br>Overall, the optimized Swin-Unet algorithm exhibits superior<br>segmentation performance and stronger **Example 12 Symmultipum**<br>
edge segmentation of green fruits, thereby effectively<br>
improving the success rate of picking in real scenarios.<br>
Overall, the optimized Swin-Unet algorithm exhibits superior<br>
segmentation perform edge segmentation of green fruits, thereby effectively improving the success rate of picking in real scenarios.<br>Overall, the optimized Swin-Unet algorithm exhibits superior segmentation performance and stronger generalizat edge segmentation of green fruits, thereby effectively improving the success rate of picking in real scenarios.<br>Overall, the optimized Swin-Unet algorithm exhibits superior segmentation performance and stronger generalizat edge segmentation of green fruits, thereby effectively<br>improving the success rate of picking in real scenarios.<br>Overall, the optimized Swin-Unet algorithm exhibits superior<br>segmentation performance and stronger generalizat improving the success rate of picking in real scenarios.<br>Overall, the optimized Swin-Unet algorithm exhibits superior<br>segmentation performance and stronger generalization ability,<br>offering a theoretical reference for the s Overall, the optimized Swin-Unet algorithm exhibits superior segmentation performance and stronger generalization ability, offering a theoretical reference for the segmentation of other green fruits. However, despite the s segmentation performance and stronger generalization ability,<br>offering a theoretical reference for the segmentation of other<br>green fruits. However, despite the significant progress made<br>by the optimized Swin-Unet algorithm offering a theoretical reference for the segmentation of other green fruits. However, despite the significant progress made by the optimized Swin-Unet algorithm compared to other semantic segmentation algorithms, its incre green fruits. However, despite the significant progress made<br>by the optimized Swin-Unet algorithm compared to other<br>semantic segmentation algorithms, its increased model<br>complexity remains a notable issue. This complexity by the optimized Swin-Unet algorithm compared to other<br>semantic segmentation algorithms, its increased model<br>complexity remains a notable issue. This complexity is<br>particularly prominent in application scenarios requiring Swin-Onet networks and exploring inote ingrived given designs. Such efforts aim to mitigate the model's dependence<br>
on computational resources without compromising<br>
segmentation accuracy, thereby enabling the optimized<br>
Sw gns. Such efforts aim to mitigate the model's dependence<br>
computational resources without compromising<br>
nentation accuracy, thereby enabling the optimized<br>
n-Unet model to better adapt to various application<br>
irements, esp computational resources without compromising<br>
aentation accuracy, thereby enabling the optimized<br>
n-Unet model to better adapt to various application<br>
irements, especially those demanding high real-time<br>
ormance and resour nentation accuracy, thereby enabling the optimized<br>
n-Unet model to better adapt to various application<br>
irements, especially those demanding high real-time<br>
ormance and resource efficiency.<br>
REFERENCES<br>
Perveen R, Suleria

## **REFERENCES**

- 919-929.<br>Tiwari J K, Singh A K, Behera T K. CRISPR/Cas genome editing in SWIN-Onet moder to better adapt to various approach<br>requirements, especially those demanding high real-time<br>performance and resource efficiency.<br><br> $REFERENCES$ <br>[1] Perveen R, Suleria H A R, Anjum F M, et al. Tomato (Solanum<br>Lycop irements, especially those demanding high real-time<br>ormance and resource efficiency.<br>
REFERENCES<br>
Perveen R, Suleria H A R, Anjum F M, et al. Tomato (Solanum<br>
Lycopersicum) carotenoids and lycopenes chemistry; metabolism,<br> STEP STEER ENCES<br>
SCIENCES<br>
REFERENCES<br>
Perveen R, Suleria H A R, Anjum F M, et al. Tomat<br>
Lycopersicum) carotenoids and lycopenes chemistry;<br>
absorption, nutrition, and allied health claims — a cor<br>
review. Critical Revie REFERENCES<br>
[1] Perveen R, Suleria H A R, Anjum F M, et al. Tomato (Solanum<br>
Lycopersicum) cartenoids and lycopenes chemistry; metabolism,<br>
absorption, nutrition, and allied health claims — a comprehensive<br>
review. Critica REFERENCES<br>
Reveen R, Suleria H A R, Anjum F M, et al. Tomato (Solanum<br>
Lycopersicum) carotenoids and lycopenes chemistry; metabolism,<br>
absorption, nutrition, and allied health claims — a comprehensive<br>
riverview. Critical REFERENCES<br>
Reven R, Suleria H A R, Anjum F M, et al. Tomato (Solanum<br>
Lycopersicum) carotenoids and lycopenes chemistry; metabolism,<br>
absorption, nutrition, and allied health claims — a comprehensive<br>
review. Critical Rev REFERENCES<br>
[1] Perveen R, Suleria H A R, Anjum F M, et al. Tomato (Solanum<br>
Lycopersicum) carotenoids and lycopenes chemistry; metabolism,<br>
absorption, nutrition, and allied health claims — a comprehensive<br>
review. Critic Perveen R, Suleria H A R, Anjum F M, et al. Tomato (Solanum<br>Lycopersicum) carotenoids and lycopenes chemistry; metabolism,<br>absorption, nutrition, and allied health claims — a comprehensive<br>review. Critical Reviews in Food [1] Cycopersicum) carotecoids and lycopenes chemistry; metabolism,<br>
Lycopersicum) carotenoids and lycopenes chemistry; metabolism,<br>
absorption, nutrition, and allied health claims — a comprehensive<br>
review. Critical Review Experistion in cunclentics and rivelopments cleanisms, included<br>absorption, nutrition, and allied health claims - a comprehensive<br>review. Critical Reviews in Food Science and Nutrition, 2015, 55(7):<br>919-929.<br>Tiwari J K, Si
- 
- 
- 
- absorption, indition, and alted ideal tradition and complements and complements of review. Critical Reviews in Food Science and Nutrition, 2015, 55(7): 919-929.<br>Tiwari J K, Singh A K, Behera T K. CRISPR/Cas genome editing From Summan Keylews in Food Science and Nutriuon, 2015, 35(7).<br>
919-929.<br>
Tiwari J K, Singh A K, Behera T K. CRISPR/Cas genome editing in<br>
tomato improvement: advances and applications. Frontiers in Plant<br>
Science, 2023, **FIST-727.**<br>Tiwari J K, Singh A K, Behera T K. CRISPR/Cas genome editing in Tiwari J K, Singh A K, Behera T K. CRISPR/Cas genome editing in Sotience, 2023, 14: 1121209.<br>Ugonna C U, Jolaoso M A, Onwualu A P. Tomato value c
- 
- Tiwari 3 K, Singit A K, Beliera 1 K. CKISFNCas genome et<br>
tomato improvement: advances and applications. Frontiers<br>
Science, 2023, 14: 1121209.<br>
Ugonna C U, Jolaoso M A, Onwualu A P. Tomato value<br>
nigeria: issues, challeng Givente, 2023, 14: 1121209.<br>
Science, 2023, 14: 1121209.<br>
13) Ugonna C U, Jolaoso M A, Onwualu A P. Tomato value chain in<br>
nigeria: issues, challenges and strategies. Journal of Scientific<br>
Research and Reports, 2015, 7(7) Suence, 2025, 14. 11-12109<br>Ugonna C U, Jolaoso M A, Onwualu A P. Tomato value chain in<br>nigeria: issues, challenges and strategies. Journal of Scientific<br>Research and Reports, 2015, 7(7): 501-515.<br>Arad B, Balendonck J, Bart Ogonial Co., Jolaso M A, Chinwalan A F. Tolliato Vatie Chain in<br>nigeria: issues, challenges and strategies. Journal of Scientific<br>nigratic issues, challenges and strategies. Journal of Scientific<br>Research and Reports, 2015 79-89.<br>Fujinaga T, Nakanishi T. Semantic segmentation of strawberry plants Fassaria and Neptons, 2015, 1, 01: 301-311.<br>
Farad B, Balendonck J, Barth R, et al. Development of a sweet pepper<br>
harvesting robot. Journal of Field Robotics, 2020, 37, 1027 - 1039.<br>
[5] Xiong Y, Ge Y, Grimstad L, et al. Arlad B, Batendonck 3, Battil R, et al. Development of a sweet pepper<br>
harvesting robot. Journal of Field Robotics, 2020, 37, 1027 – 1039.<br>
Xiong Y, Ge Y, Grimstad L, et al. An autonomous strawberry<br>
harvesting robot: desi narvesung root. Journal of Field Robotics, 2020, 57, 1027 = 1059.<br>
Xiong Y, Ge Y, Grimstad L, et al. An autonomous strawberry -<br>
harvesting robot: design, development, integration, and field<br>
evaluation. Journal of Field R
- 
- [5] Xiong Y, Ge Y, Grimstad L, et al. An autonomous strawberry -<br>harvesting robot: design, development, integration, and field<br>evaluation. Journal of Field Robotics, 2020, 37(2): 202-224.<br>[6] Jia W, Zhang Y, Lian J, et al. harvesting robot: design, development, integration, and field evaluation. Journal of Field Robotics, 2020, 37(2): 202-224.<br>Jia W, Zhang Y, Lian J, et al. Apple harvesting robot under information dichnology: A review. Inter evaluation. Journal of Field Robotics, 2020, 37(2): 202-224.<br>Jia W, Zhang Y, Lian J, et al. Apple harvesting robot under information<br>technology: A review. International Journal of Advanced Robotic<br>Eystems, 2020, 17(3): 253 Jia W, Zhang Y, Lian J, et al. Apple harvesting robot under information<br>technology: A review. International Journal of Advanced Robotic<br>Systems, 2020, 17(3): 25310.<br>Saranya N, Srinivasan K, Pravin Kumar S K, et al. Fruit c 2023.<br>
[10] Häni N, Roy P, Isler V. A comparative study of fruit detection and Systems, 2020, 17(3): 25310.<br>
[7] Saranya N, Srinivasan K, Pravin Kumar S K, et al. Fruit classification<br>
using traditional machine learning and deep learning approach.<br>
Computational Vision and Bio-Inspired Computing: ICC Saranya N, Srinivasan K, Pravin Kumar S K, et al. Fruit classification<br>
using traditional machine learning and deep learning approach.<br>
Computational Vision and Bio-Inspired Computing: ICCVBIC, 2020:<br>
79-89.<br>
Fujinaga T, N using traditional machine learning and deep learning a<br>Computational Vision and Bio-Inspired Computing: ICCVBI<br>79-89.<br>Fujinaga T, Nakanishi T. Semantic segmentation of strawber<br>using DeepLabV3+ for small agricultural robot Computational Vision and Bio-Inspired Computing: ICCVBIC, 2020:<br>
79-89.<br>
[8] Fujinaga T, Nakanishi T. Semantic segmentation of strawberry plants<br>
using DeepLabV3+ for small agricultural robot. International<br>
Symposium on S 79-89.<br>
Fujinga T, Nakanishi T. Semantic segmentation of strawberry plants<br>
using DeepLabV3+ for small agricultural robot. International<br>
Symposium on System Integration. IEEE, 2023: 1-6.<br>
Marizuana Mat Daud, Zulaikha Kadi Fujinaga T, Nakanishi T. Semantic segmenta<br>
using DeepLabV3+ for small agricultiu<br>
Symposium on System Integration. IEEE, 20<br>
Marizuana Mat Daud, Zulaikha Kadim,<br>
Detection of oil palm tree and loose fruitle<br>
ready-to-harv using DeepLabV3+ for small agricultural robot. International<br>
Symposium on System Integration. IEEE, 2023: 1-6.<br>
IPM Marizuana Mat Daud, Zulaikha Kadim, and Hon Hock Woon,<br>
Detection of oil palm tree and loose fruitlets fo Symposium on System Integration. IEEE, 2023: 1-6.<br>
Marizuana Mat Daud, Zulaikha Kadim, and Hon Hock Woon,<br>
Detection of oil palm tree and loose fruites for fresh fruit bunch's<br>
ready-to-harvest prediction via deep learning Marizuana Mat Daud, Zulaikha Kadım, and Hon Hock Woon,<br>Detection of oil palm tree and loose fruitlets for fresh fruit bunch's<br>ready-to-harvest prediction via deep learning approach. IAENG<br>International Journal of Computer Detection of oil palm tree and loose fruitlets for fresh fruit bunch's<br>
ready-to-harvest prediction via deep learning approach. IAENG<br>
International Journal of Computer Science, vol. 50, no.4, pp1183-1193,<br>
2023.<br>
[10] Hän
- ready-to-harvest prediction via deep learning approach. IAENG<br>International Journal of Computer Science, vol. 50, no.4, pp1183-1193,<br>2023.<br>Häni N, Roy P, Isler V. A comparative study of fruit detection and<br>counting methods International Journal of Computer Science, vol. 50, no.4, pp112<br>2023.<br>Häni N, Roy P, Isler V. A comparative study of fruit detect<br>counting methods for yield mapping in apple orchards. Journal<br>Robotics, 2020, 37(2): 263-282 2023.<br>
[10] Häni N, Roy P, Isler V. A comparative study of fruit detection and<br>
counting methods for yield mapping in apple orchards. Journal of Field<br>
Robotics, 2020, 37(2): 263-282.<br>
[11] Bargoti S, Underwood J. Deep fru
- Häni N, Roy P, Isler V. A comparative study of fruit detection and<br>counting methods for yield mapping in apple orchards. Journal of Field<br>Robotics, 2020, 37(2): 263-282.<br>Bargoti S, Underwood J. Deep fruit detection in orch
- 
- 
- 
- counting methods for yield mapping in applie orchards<br>
Robotics, 2020, 37(2): 263-282.<br>
Bargoti S, Underwood J. Deep fruit detection in orcl<br>
International Conference on Robotics and Automatic<br>
International Conference on Robotics, 2020, 37(2): 263-282.<br>
[11] Bargoti S, Underwood J. Deep fruit detection in orchards. 2017 IEEE<br>
International Conference on Robotics and Automation (ICRA). IEEE,<br>
2017: 3626-3633.<br>
[12] Barth R, IJsselmuiden J, Bargoti S, Underwood J. Deep fruit detection in orchards. 2017 IEEE<br>International Conference on Robotics and Automation (ICRA). IEEE,<br>2017: 3626-3633.<br>Barth R, IJsselmuiden J, Hemming J, et al. Synthetic bootstrapping of<br>c International Conference on Robotics and Automation (ICRA). IEEE,<br>2017: 3626-3633.<br>Barth R, IJsselmuiden J, Hemming J, et al. Synthetic bootstrapping of<br>Barth R, IJsselmuiden J, Hemming J, et al. Synthetic bootstrapping of 2017: 3626-3633.<br>Barth R, IJsselmuiden J, Hemming J, et al. Synth<br>convolutional neural networks for semantic plan<br>Computers and Electronics in Agriculture, 2019, 1<br>Kang H, Chen C. Fruit Detection, Segmentation and<br>of envir
- **Engineering Letters**<br>
[16] Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365-
- **Engineering Letters**<br>
Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365-3375. **Engineering Letters**<br>Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>industrial framework for rotten and fresh fruit detection using semantic<br>segmentation. Microsystem Technologies, 2021, 27: 3365-3375.<br>Li **Engineering Letters**<br>
[16] Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365-**Engineering Letters**<br>
Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365-3375. **Engineering Letters**<br>
Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365-3375. **Engineering Letters**<br>
[16] Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365-**Engineering Letters**<br>
Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365-3375. **Engineering Letters**<br>
[16] Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365
- 
- 
- Engineer ing Ecclesis<br>
Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
esgementation. Microsystem Technologies, 2021, 27: 3365-3375 [16] Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>
industrial framework for rotten and fresh fruit detection using semantic<br>
segmentation. Microsystem Technologies, 2021, 27: 3365-3375.<br>
[17] Li Q, Jia Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>industrial framework for rotten and fresh fruit detection using semantic<br>segmentation. Microsystem Technologies, 2021, 27: 3365-3375.<br>Li Q, Jia W, Sun M, et a Roy K, Chaudhuri S S, Pramanik S. Deep learning based real-time<br>industrial framework for rotten and fresh fruit detection using semantic<br>segmentation. Microsystem Technologies, 2021, 27: 3365-3375.<br>Li Q, Jia W, Sun M, et a Roy K, Chaudhuri S S, Pramanik S. Deep learning based reindustrial framework for rotten and fresh fruit detection using segmentation. Microsystem Technologies, 2021, 27: 3365-337: Li Q, Jia W, Sun M, et al. A novel green a molustrial framework for rotten and fresh fruit detection using sementiation. Microsystem Technologies, 2021, 27: 3365-3375.<br>
[17] Li Q, Jia W, Sun M, et al. A novel green apple segmentation algorithm<br>
based on ensemble Usegmentation. Microsystem Technologies, 2021, 27: 3365-3375.<br>
Li Q, Jia W, Sun M, et al. A novel green apple segmentation algorithm<br>
based on ensemble U-Net under complex orchard environment.<br>
Computers and Electronics in Li Q, Jia W, Sun M, et al. A novel green app<br>based on ensemble U-Net under compl<br>Computers and Electronics in Agriculture, 2<br>He J, Duan J, Yang Z, et al. Method for segn<br>based on improved DeepLabv3+. Agronomy<br>Yan C, Chen Z based on ensemble U-Net under complex orchard environment.<br>
Computers and Electronics in Agriculture, 2021, 180: 105900.<br>
[18] He J, Duan J, Yang Z, et al. Method for segmentation of banana crown<br>
based on improved DeepLa Computers and Electronics in Agriculture, 2021, 180: 105900.<br>He J, Duan J, Yang Z, et al. Methof for segmentation of banana crown<br>based on improved DeepLabv3+. Agronomy, 2023, 13(7): 1838.<br>Yan C, Chen Z, Li Z, et al. Tea s He J, Duan J, Yang Z, et al. Method for segmentation<br>based on improved DeepLabv3+. Agronomy, 2023,<br>Yan C, Chen Z, Li Z, et al. Tea sprout picking po<br>based on improved DeepLabV3+. Agriculture, 2022,<br>Bai Y, Guo Y, Zhang Q, e based on improved DeepLabv3+. Agronomy, 2023, 13(7): 1838.<br>
[19] Yan C, Chen Z, Li Z, et al. Tea sprout picking point identification<br>
based on improved DeepLabV3+. Agriculture, 2022, 12(10): 1594.<br>
[20] Bai Y, Guo Y, Zhang Yan C, Chen Z, Li Z, et al. Tea sprout picking point identification<br>based on improved DeepLabV3+. Agriculture, 2022, 12(10): 1594.<br>Bai Y, Guo Y, Zhang Q, et al. Multi-network fusion algorithm with<br>transfer learning for gr based on mproved DeepLabV3+. Agriculture, 2022, 12(10): 1594.<br>Bai Y, Guo Y, Zhang Q, et al. Multi-network fusion algorithm w<br>transfer learning for green cucumber segmentation and recognit<br>under complex natural environment. [20] Ban Y, Giuo Y, Zhang Q, et al. Multi-network fusion algorithm with<br>transfer learning for green cucumber segmentation and recognition<br>under complex natural environment. Computers and Electronics in<br>Agriculture, 2022, 1
- under complex natural environment. Computers and I<br>Agriculture, 2022, 194: 106789.<br>Liu X, Yu J, Kurihara T, et al. Hyperspectral imaging for<br>segmentation using a complex-valued neural network.<br>265: 169527.<br>Cao H, Wang Y, C
- 
- 
- transter learning for green cucumber segmentation and recognition<br>
ander complex natural environment. Computers and Electronics in<br>
Agriculture, 2022, 194: 106789.<br>
Liu X, Yu J, Kurihara T, et al. Hyperspectral imaging for Agriculture, 2022, 194: 106789.<br>
[21] Liu X, Yu J, Kurihara T, et al. Hyperspectral imaging for green pepper<br>
segmentation using a complex-valued neural network. Optik, 2022,<br>
265: 169527.<br>
[22] Cao H, Wang Y, Chen J, et Liu X, Yu J, Kurihara T, et al. Hyperspectral imaging for green pepper<br>segmentation using a complex-valued neural network. Optik, 2022,<br>265: 169527.<br>Cao H, Wang Y, Chen J, et al. Swin-unet: Unet-like pure transformer<br>for m
- 
- segmentation using a complex-valued neural netw<br>265: 169527.<br>Cao H, Wang Y, Chen J, et al. Swin-unet: Unet-lik<br>for medical image segmentation. European confer<br>vision, 2022: 205-218.<br>Russell B C, Torralba A, Murphy K P, et 265: 169527.<br>
221 Cao H, Wang Y, Chen J, et al. Swin-unet: Unet-like pure transformer<br>
for medical image segmentation. European conference on computer<br>
vision, 2022: 205-218.<br>
23] Russell B C, Torralba A, Murphy K P, et al Cao H, Wang Y, Chen J, et al. Swin-unet: Unet-like pure transformer<br>for medical image segmentation. European conference on computer<br>vision, 2022: 205-218.<br>Russell B C, Torralba A, Murphy K P, et al. LabelMe: a database and for medical image segmentation. European conference on computer<br>
Nussell B C, Torralba A, Murphy K P, et al. LabelMe: a database and<br>
web-based tool for image annotation. International Journal of<br>
Computer Vision, 2008, 77 vision, 2022: 205-218.<br>
Russell B C, Torralba A, Murphy K P, et al. LabelMe: a database and<br>
web-based tool for image annotation. International Journal of<br>
Computer Vision, 2008, 77: 157-173.<br>
Lin T Y, Maire M, Belongie S, [23] Russell B C, Iorralba A, Murphy K P, et al. LabelMe: a database and<br>
web-based tool for image annotation. International Journal of<br>
Computer Vision, 2008, 77: 157-173.<br>
[24] Lin T Y, Maire M, Belongie S, et al. Micros web-based tool for image annotation. International Journal of<br>Computer Vision, 2008, 77: 157-173.<br>Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common<br>objects in context[C]. European Conference on Computer Vision<br>Zu Computer Vision, 2008, 77: 157-173.<br>
[24] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common<br>
objects in context(C]. European Conference on Computer Vision<br>
Zurich: Springer, 2014.<br>
[25] Zhang W, Zhao Y, Guan Y, e Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common<br>objects in context[C]. European Conference on Computer Vision<br>Zurich: Springer, 2014.<br>Zhang W, Zhao Y, Guan Y, et al. Green apple detection method based<br>on optimi blyects in context[C]. European Conference on Computer Vision<br>
Zinang W, Zhan Y, Guan Y, et al. Green apple detection method based<br>
2023, 31(3): 1104-1113.<br>
Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: Semantic ima Zurich: Springer, 2014.<br>
[25] Zhang W, Zhao Y, Guan Y, et al. Green apple detection method based<br>
on optimized yolov5 under orchard environment. Engineering Letters,<br>
2023, 31(3): 1104-1113.<br>
[26] Chen L C, Papandreou G, Zhang W, Zhao Y, Guan Y, et al. Green apple detection method based<br>
2023, 31(3): 1104-1113.<br>
2023, 31(3): 1104-1113.<br>
Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: Semantic image<br>
segmentation with deep convolutional on optimized yolov5 under orchard environment. Engineering Letters,<br>2023, 31(3): 1104-1113.<br>Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: Semantic image<br>esgmentation with deep convolutional nets, atrous convolution, 2023, 31(3): 1104-1113.<br>
[26] Chen L C, Papandrou G, Kokkinos I, et al. Deeplab: Semantic image<br>
segmentation with deep convolutional nets, atrous convolution, and<br>
fully connected crfs. IEEE Transactions on Pattern Analy
- 
- 
- 
- segmentation with deep convolutional nets, atrous convolution, and<br>fully connected crfs. IEEE Transactions on Pattern Analysis and<br>Machine Intelligence, 2017, 40(4): 834-848.<br>Oktay O, Schlemper J, Folgoc L L, et al. Attent 1-16.<br>
[31] Chen L C. Zhu Y. Papandreou G. et al. Encoder-decoder with atrous Machine Intelligence, 2017, 40(4): 834-848.<br>
[27] Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: Learning<br>
where to look for the pancreas. arXiv preprint arXiv:1804.03999, 2018.<br>
[28] Liu R, Tao F, Liu X, et al. Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: Learning<br>where to look for the pancreas. arXiv preprint arXiv: 1804.03999, 2018.<br>Liu R, Tao F, Liu X, et al. RAANet: a residual ASPP with attention<br>framework for se where to look for the pancreas. arXiv preprint arXiv:1804.03999, 2018.<br>
Liu R, Tao F, Liu X, et al. RAANet: a residual ASPP with attention<br>
framework for semantic segmentation of high-resolution remote<br>
sensing images. Rem framework for semantic segmentation of high-resolution remote<br>
29) Zhang J, Qin Q, Ye Q, et al. ST-unet: Swin transformer boosted U-net<br>
with cross-layer feature enhancement for medical image segmentation.<br>
Computers in Bi sensing images. Remote Sensing, 2022, 14(13): 3109.<br>
Zhang J, Qin Q, Ye Q, et al. ST-unet: Swin transformer boosted U-net<br>
with cross-layer feature enhancement for medical image segmentation.<br>
Computers in Biology and Medi Zhang J, Qm Q, Ye Q, et al. ST-unet: Swin transformer both cross-layer feature enhancement for medical image se<br>Computers in Biology and Medicine, 2023, 153: 106516.<br>Zheng Z, Liang E, Zhang Y, et al. A segmentation-based a with cross-layer feature enhancement for medical image segmentation.<br>
Computers in Biology and Medicine, 2023, 153: 106516.<br>
[30] Zheng Z, Liang E, Zhang Y, et al. A segmentation-based algorithm for<br>
classification of beni Computers in Biology and Medicine, 2023, 153: 106516.<br>
Zheng Z, Liang E, Zhang Y, et al. A segmentation-based algorithm for<br>
classification of benign and malignancy thyroid nodules with<br>
multi-feature information. Biomedic
- 801-818. Zheng Z, Liang E, Zhang Y, et al. A segmentation-based al<br>classification of benign and malignancy thyroid no<br>multi-feature information. Biomedical Engineering Let<br>1-16.<br>1-16.<br>1-16.<br>1-16.<br>1-16.<br>1-16.<br>1-16.<br>1-16.<br>1-16.<br>1-16. classification of benign and malignancy thyroid nodules with<br>multi-feature information. Biomedical Engineering Letters, 2024:<br>1-16.<br>131] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous<br>separable convoluti multi-feature information. Biomedical Engineering Letters, 2024:<br>1-16.<br>Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous<br>separable convolution for semantic image segmentation. Proceedings<br>601-818.<br>Chen L C, 1-16.<br>Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous<br>separable convolution for semantic image segmentation. Proceedings<br>of the European Conference on Computer Vision (ECCV). 2018:<br>801-818.<br>Chen L C, Papa [31] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous<br>separable convolution for semantic image segmentation. Proceedings<br>of the European Conference on Computer Vision (ECCV). 2018:<br>801-818.<br>801-818.<br>601-8
- of the European Conference on Computer Vi<br>
801-818.<br>
Chen L C, Papandreou G, Schroff F, et al<br>
convolution for semantic image segmentat<br>
arXiv:1706.05587, 2017.<br>
Zhao H, Shi J, Qi X, et al. Pyramid Scen<br>
Proceedings of the
- 
- separable convolution for semantic image segmentation. Proceedings<br>801-818.<br>801-818.<br>Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous<br>convolution for semantic image segmentation. arXiv preprint<br>arXiv:1706.05587, 801-818.<br>
[32] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous<br>
convolution for semantic image segmentation. arXiv preprint<br>
arXiv:1706.05587, 2017.<br>
[33] Zhao H, Shi J, Qi X, et al. Pyramid Scene Parsing Netwo Zhao H, Shi J, Qi X, et al. Pyramid Scene Parsin<br>Proceedings of the IEEE Conference on Computer Vision<br>Recognition. 2017: 2881-2890.<br>Fu J, Liu J, Tian H, et al. Dual attention network<br>Segmentation. Proceedings of the IEEE/
- Fu J, Liu J, Tian H, et al. Dual attention<br>segmentation. Proceedings of the IEEE/CVF Cc<br>Vision and Pattern Recognition. 2019: 3146-31:<br>Zhang W, Pang J, Chen K, et al. K-net: To<br>segmentation. Advances in Neural Information<br>
- 
- Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous<br>convolution for semantic image segmentation. arXiv preprint<br>arXiv:1706.05587, 2017.<br>Zhao H, Shi J, Qi X, et al. Pyramid Scene Parsing Network.<br>Proceedings of the convolution for semantic image segmentation. arXiv preprint<br>
arXiv: 1706.05587, 2017.<br>
[33] Zhao H, Shi J, Qi X, et al. Pyramid Scene Parsing Network.<br>
Proceedings of the IEEE Conference on Computer Vision and Pattern<br>
Rec arXiv:1706.05587, 2017.<br>
Zhao H, Shi J, Qi X, et al. Pyramid Scene Parsing Network.<br>
Zhao H, Shi J, Qi X, et al. Dual attention network for scene<br>
Recognition. 2017: 2881-2890.<br>
Fu J, Liu J, Tian H, et al. Dual attention n Proceedings of the IEEE Contrence on Computer Vision and Pattern<br>
Recognition. 2017: 2881-2890.<br>
[34] Fu J, Liu J, Tian H, et al. Dual attention network for scene<br>
segmentation. Proceedings of the IEEE/CVF Conference on Co Recognition. 2017: 2881-2890.<br>Fu J, Liu J, Tian H, et al. Dual attention network for scene<br>segmentation. Proceedings of the IEEE/CVF Conference on Computer<br>Vision and Pattern Recognition. 2019: 3146-3154.<br>Zhang W, Pang J, segmentation. Proceedings of the IEEE/CVF Conference on Computer<br>
1951 Zhang W, Pang J, Chen K, et al. K-net: Towards unified image<br>
segmentation. Advances in Neural Information Processing Systems,<br>
2021, 34: 10326-10338.<br> Vision and Pattern Recognition. 2019: 3146-3154.<br>
Zhang W, Pang J, Chen K, et al. K-net: Towards unified image<br>
segmentation. Advances in Neural Information Processing Systems,<br>
2021, 34: 10326-10338.<br>
Huang L, Yuan Y, Guo
- 
-