# *A* **Semantic SLAM Integrated with Enhanced**<br> *YOLOv7 Target Detection Algorithm*<br> *ZhangFang Hu, FangYu Li, JiXiang Shen*<br> *Abstract*—This paper proposes a semantic SLAM integrated relies on understanding the environment A Semantic SLAM Integrated with<br>YOLOv7 Target Detection Algor<br>ZhangFang Hu, FangYu Li, JiXiang Shen<br>ZhangFang Hu, FangYu Li, JiXiang Shen<br>Matract—This paper proposes a semantic SLAM integrated<br>with an enhanced YOLOv7 targe Engineering Letters<br>
A Semantic SLAM Integrated with Enhanced<br>
YOLOv7 Target Detection Algorithm<br>
ZhangFang Hu, FangYu Li, JiXiang Shen<br>
real-time and robust. Traditional visual SLAM primarily<br>
real-time and robust. Tradit Engineering Letters<br>
emantic SLAM Integrated with Enhanced<br>
YOLOv7 Target Detection Algorithm<br>
ZhangFang Hu, FangYu Li, JiXiang Shen<br>
real-time and robust. Traditional visual SLAM primarily<br>
real-time and robust. Tradition Engineering Letters<br>
SLAM Integrated with Enhanced<br>
Target Detection Algorithm<br>
ZhangFang Hu, FangYu Li, JiXiang Shen<br>
real-time and robust. Traditional visual SLAM primar<br>
real-time and robust. Traditional visual SLAM pri

**address the issue of image blurring caused by robot movement**<br> **address the issue of image blurring caused by robot movement**<br> **address the issue of image blurring caused by robot movement**<br> **address the issue of image bl and camera shake, we have incorporated an image incorporation of the resulting images are more clearer. In the feature shake, we have incorporated an image enhancement module before the tracking thread. Consequently, the EXECTION / LATGET DETECTION**<br>
ZhangFang Hu, FangYu Li, JiXiang SI<br> *Abstract*—This paper proposes a semantic SLAM integrated<br>
with an enhanced YOLOv7 target detection algorithm. To<br>
address the issue of image blurring cau **Example 19 and Consequently,** The resulting images are more clearer. In the feature of the resulting images are more clearer. In the feature of image by the resulting cause of image by root movement of the and camera sha ZhangFang Hu, FangYu Li, JiXiang Shen<br>real-time and robust.<br>Mintegrated relies on understanding<br>with an enhanced YOLOv7 target detection algorithm. To<br>features of images [6],<br>address the issue of image blurring caused by r **ChangFang Hu, FangYu Li, JiXiang Shen**<br> **cal-time and robust.**<br> **cal-time and robust**<br> **cal-time and robust**<br> **cal-time and robust**<br> **cal-time and robust**<br> **cal-time and robustive in the influence**<br> **cal-time and robustiv** ZhangFang Hu, FangYu Li, JiXiang Shen<br> *Abstract*—This paper proposes a semantic SLAM integrated<br>
with an enhanced YOLOv7 target detection algorithm. To<br>
address the issue of image blurring caused by robot movement<br>
and ca *POLORY* **Abstract—This paper proposes a semantic SLAM integrated** relies on understanding the with an enhanced YOLOv7 target detection algorithm. To features of images [6], which address the issue of image blurring caused *i***m and robust.<br>
<b>integrated integrated i** cal-time and robust.<br> **integrated i** relies on understanding<br> **integrated i** relies on understanding<br> **integrated i** relies of **images** lurring caused by robus movemen *followith* **an enhanced YOLOv7 target detection algorithm. To with an enhanced YOLOv7 target detection algorithm. To features of images [6], whice and camera shake, we have incorporated an image burring cause it only foc Abstract—This paper proposes a semantic SLAM integrated**<br>
with an enhanced YOLOv7 target detection algorithm. To<br>
seal cannot cannot and carres of images [6], whis<br>
and camera shake, we have incorporated an image<br>
enhance **Provided and the system in Super provides a sensation Example and canners are the solutions of images [6],** and canners hake, we have incorporated an image invironment. Tradition enhancement module before the tracking thr What are unauthed TOLOW arget detection and the similar semantic SLAM.<br>
Indees the correct of the compared to the care of image burring caused by robot movement because it only focuses on geonal<br>
and camera shake, we have and cameas ue issue of image butting caused by root invertint. Traditional VS<br>and camera shake, we have incorporated an image environment. Traditional VS<br>the resulting images are more clearer. In the feature extraction<br>can and callera shake, we have incorporated an<br>enhancement module before the tracking thread. Consect<br>the resulting images are more clearer. In the feature ext<br>stage, we introduce adaptive thresholds to improve the s<br>capabilit **INTEONICTION**<br> **INTEONICTION**<br> **INSPECTED:** INCREDUCTION<br> **INSPECTED:** INCREDUCTION<br> **IDENTIFY AND SECUTE TO THE CONSEQUENT**<br> **INSPECTED:** INCREDUCTION<br> **IDENTIFY AND SECUTE IN THE CONSEQUENT**<br> **INSPECTED:** THE CONSEQUENT **Example 19 and the comparison of dynamic displane capability in feature point extraction. To minimize the influence** automob<br> **adaptive thresholds** on this system, we employ an enhanced<br> **POLOv7** algorithm to detect dynam It is system, we employ an enhanced<br>to detect dynamic targets. Then, we<br>ipolar constraint to eliminate dynamic<br>dy, We evaluated our system with five<br>the TUM dataset, and compared with<br>tem improves more than 91% in accuracy grate it with epipolar constraint to eliminate dynamic<br>
urge points. Finally, We evaluated our system with five<br>
ences taken from the TUM dataset, and compared with<br>
ences taken from the TUM dataset, and compared with<br>
the

refers to a process where a robot, without any prior<br>
references than from the TUM dataset, and compared with five<br>
in B-SLAM3, our system improves more than 91% in accuracy,<br>
capability to comprehend the de-<br>
to 98%. More sequences taken from the TUM dataset, and one pared with<br>
ORB-SLAM3, our system improves more than 91% in accuracy, capability to comprehend the environment<br>
or to 98%. Moreover, compared to similar semantic SLAM fails to ORESAMIN, our system improves more than 91% in accuracy, capacing to 98%. Moreover, compared to similar semantic SLAM fails to meet the demands for systems, our system offers improved accuracy as well as the current era of used, researchers divide SLAM into Laser SLAM (LS) and<br>
visual SLAM (WS). Among them, VS has the benefits of traditional visual<br>
visual SLAM into Laser SLAM into Laser SLAM and a prevalence of the entimies of traditional v Finance of the terms and many and the terms of traditional visual states of minimize the performance.<br> *Interns* and the performance SLAM, image enhancement, designed for dynamic environments, the tem dynamic environments, *Index Terms*—semantic SLAM, image enhancement,<br>
adaptive thresholds, YOLOv7, epipolar constraint<br>
in dynamic environments, the<br>
to minimize the influence of more system by excluding feature<br>
position of more system by exc *Index Terms*—semantic SLAM, image enhancement,<br>adaptive thresholds, YOLOv7, epipolar constraint<br>to minimize the influence<br>system by excluding feat<br>I. INTRODUCTION<br>surrounding feat<br>objects. This challenge is<br>expectinc-ba<br>o adaptive thresholds, YOLOv7, epipolar constraint<br>
to minimize the influ<br>
system by excluding<br>
Sy I. INTRODUCTION<br>
I. INTRODUCTION<br>
System by excluding feature<br>
objects. This challenge is<br>
objects. This challenge is<br>
objects. This challenge is<br>
approaches: geometric-base<br>
information, simultaneously localizes itself an I. INTRODUCTION<br>
System by exaction<br>
objects. This challenge<br>
information, simultaneously localization and Mapping (SLAM)<br>
information, simultaneously localizes itself and constructs a<br>
map of its surrounding area[1]. Base 1. INTRODUCTION<br>
1. INTRODUCTION<br>
1. INTRODUCTION<br>
information, simultaneously localizes itself and constructs a<br>
green information, simultaneously localizes itself and constructs a<br>
deep<br>
used, researchers divide SLAM int **IMULTANEOUS** Localization and Mapping (SLAM) applycaties. geometry-based dynamic volume of its surrounding areal [1]. Based on the ypes of sensors of sensors of surrounding and constructs and constructs and specific tech sual SLAM (VS). Among them, VS has the benefits of<br>
duced expenses and access to obtain more data from<br>
rroundings, which can give mobile robots stronger<br>
rroundings, which can give mobile robots stronger<br>
experiment, such From Sucheme, such as the Rand<br>
in the detect expenses and access to obtain more data from<br>
informulation (RANSAC [8]). In additive<br>
introducing supported in particular political replications, visual SLAM needs to be<br>
Manu SIMULTANEOUS Localization and Mapping (SLAM) approaches:<br>refers to a process where a robot, without any prior<br>specific te

Engine Dasse Stroke and Formular Corant No. Cstc2017jeyjAX0212), and the Science and Technology filtering of Research Program of Chongqing Municipal Education Commission (KJ1704072).<br>
Zhangfang Hu is a Professor at the Key Chonging, 400065, China (e-mail: s2204310665, China (e-mail: s220432006@stu.cqupt.edu.cn)<br>
Chongqing, 400065, China (e-mail: 3565207151@qq.com)<br>
Frangyn Li is a graduate student of the School of Optoelectronic Engineering,

real-time and robust. Traditional visual SLAM primarily<br>ru Li, JiXiang Shen<br>real-time and robust. Traditional visual SLAM primarily<br>relies on understanding the environment through geometric<br>features of images [6], which ha relies on **Michan Endine Conder**<br> **relies on Michan Algorithm**<br> **relies on understanding the environment through geometric**<br> **relies on understanding the environment through geometric**<br> **features of images** [6], which has Summing the Control of the Discreed Stection Algorithm<br>
Full, JiXiang Shen<br>
Full, JiXiang Shen<br>
Full is shared and robust. Traditional visual SLAM primarily<br>
relies on understanding the environment through geometric<br>
featu Example 1920 With Enhanced<br>
Stection Algorithm<br>
Tu Li, JiXiang Shen<br>
real-time and robust. Traditional visual SLAM primarily<br>
relies on understanding the environment through geometric<br>
features of images [6], which has hig Exact WITH ETHIMITS<br>
Exact WITH ETHIMITS<br>
The CONSTRIMENT CONSTRIMENT CULTURE (The and robust. Traditional visual SLAM primarily<br>
relies on understanding the environment through geometric<br>
features of images [6], which has **Example 18 Stationary School Algorithm**<br>
Stationary Shen<br>
For Li, JiXiang Shen<br>
For an and robust. Traditional visual SLAM primarily<br>
relies on understanding the environment through geometric<br>
features of images [6], whic FICCIIOII AIGOI IUIIII<br>real-time and robust. Traditional visual SLAM primarily<br>relies on understanding the environment through geometric<br>features of images [6], which has high real-time performance<br>because it only focuses The U.S. In Sixtem and The authorius State the sentence of magnetic scattering fields on understanding the environment through geometric features of images [6], which has high real-time performance because it only focuses The Valuety Comparison of the accuracy of VS systems.<br>Consequently relies on understanding the environment through geometric features of images [6], which has high real-time performance because it only focuses on geometric The U. F. JiXiang Shen<br>
real-time and robust. Traditional visual SLAM primarily<br>
relies on understanding the environment through geometric<br>
features of images [6], which has high real-time performance<br>
because it only focu real-time and robust. Traditional visual SLAM primarily<br>relies on understanding the environment through geometric<br>features of images [6], which has high real-time performance<br>because it only focuses on geometric features i real-time and robust. Traditional visual SLAM primarily<br>relies on understanding the environment through geometric<br>features of images [6], which has high real-time performance<br>because it only focuses on geometric features i real-time and robust. Traditional visual SLAM primarily<br>relies on understanding the environment through geometric<br>features of images [6], which has high real-time performance<br>because it only focuses on geometric features i relies on understanding the environment through geometric<br>features of images [6], which has high real-time performance<br>because it only focuses on geometric features in the<br>environment. Traditional VS relies on the assumpti features of images [6], which has high real-time performance<br>because it only focuses on geometric features in the<br>environment. Traditional VS relies on the assumption of a<br>stationary surroundings. But it doesn't hold true because it only focuses on geometric features in the environment. Traditional VS relies on the assumption of a stationary surroundings. But it doesn't hold true in real-world scenarios where moving entities, such as walker environment. Traditional VS relies on the assumption of a<br>stationary surroundings. But it doesn't hold true in real-world<br>scenarios where moving entities, such as walkers and<br>automobiles, are unavoidably present. Dynamic<br>e it doesn't hold true in real-world<br>
enarios where moving entities, such as walkers and<br>
tomobiles, are unavoidably present. Dynamic<br>
vironments generate numerous incorrect data associations<br>
J, leading to a reduction in th scenarios where moving entities, such as walkers and<br>automobiles, are unavoidably present. Dynamic<br>environments generate numerous incorrect data associations<br>[7], leading to a reduction in the accuracy of VS systems.<br>Conse automobiles, are unavoidably present. Dynamic<br>environments generate numerous incorrect data associations<br>[7], leading to a reduction in the accuracy of VS systems.<br>Consequently, traditional VS systems exhibit lower<br>robustn environments generate numerous incorrect data associations [7], leading to a reduction in the accuracy of VS systems.<br>Consequently, traditional VS systems exhibit lower robustness. Furthermore, traditional visual SLAM lack

Featured expenses and access to obtain finde data from the surroundings, which can give mobile robots stronger (RANSAC [8]). In additimulti-sensor fusion to localization and map building techniques have been widely environ surroundings, which can give mobile robots stronger<br>
environmental awareness [2]. So, vision-based simultaneous<br>
studied and applied to robot navigation [3], unmanned<br>
driving [4] and virtual reality [5].<br>
In practical app [7], leading to a reduction in the accuracy of VS systems.<br>Consequently, traditional VS systems exhibit lower<br>robustness. Furthermore, traditional visual SLAM lacks the<br>capability to comprehend the environment at a high le Consequently, traditional VS systems exhibit lower<br>robustness. Furthermore, traditional visual SLAM lacks the<br>capability to comprehend the environment at a high level and<br>fails to meet the demands for human-computer intera robustness. Furthermore, traditional visual SLAM lacks the capability to comprehend the environment at a high level and fails to meet the demands for human-computer interaction in the current era of intelligent technology. capability to comprehend the environment at a high level and<br>fails to meet the demands for human-computer interaction in<br>the current era of intelligent technology. To address the<br>limitations of traditional visual SLAM, vis fails to meet the demands for human-computer interaction in<br>the current era of intelligent technology. To address the<br>limitations of traditional visual SLAM, visual SLAM<br>designed for dynamic environments has emerged.<br>In dy the current era of intelligent technology. To address the limitations of traditional visual SLAM, visual SLAM designed for dynamic environments has emerged.<br>
In dynamic environments, the primary objective of VS is to minim limitations of traditional visual SLAM, visual SLAM<br>designed for dynamic environments has emerged.<br>In dynamic environments, the primary objective of VS is<br>to minimize the influence of moving entities on the SLAM<br>system by designed for dynamic environments has emerged.<br>
In dynamic environments, the primary objective of VS is<br>
to minimize the influence of moving entities on the SLAM<br>
system by excluding feature points associated with these<br>
o In dynamic environments, the primary objective of VS is<br>to minimize the influence of moving entities on the SLAM<br>system by excluding feature points associated with these<br>objects. This challenge is addressed through two dis to minimize the influence of moving entities on the SLAM<br>system by excluding feature points associated with these<br>objects. This challenge is addressed through two distinct<br>approaches: geometric-based dynamic visual SLAM, e system by excluding feature points associated with these<br>objects. This challenge is addressed through two distinct<br>approaches: geometric-based dynamic visual SLAM and<br>deep learning-based dynamic visual SLAM, each utilizing objects. This challenge is addressed through two distinct approaches: geometric-based dynamic visual SLAM and deep learning-based dynamic visual SLAM, each utilizing specific techniques designed for this purpose. Geometryapproaches: geometric-based dynamic visual SLAM and<br>deep learning-based dynamic visual SLAM, each utilizing<br>specific techniques designed for this purpose.<br>Geometry-based dynamic visual SLAM employs geometric<br>information of deep learning-based dynamic visual SLAM, each utilizing<br>specific techniques designed for this purpose.<br>Geometry-based dynamic visual SLAM employs geometric<br>information of the environment to eliminate dynamic features,<br>and specific techniques designed for this purpose.<br>Geometry-based dynamic visual SLAM employs geometric<br>information of the environment to eliminate dynamic features,<br>and a prevalent approach is the maximum consistency<br>scheme, Geometry-based dynamic visual SLAM employs geometric<br>information of the environment to eliminate dynamic features,<br>and a prevalent approach is the maximum consistency<br>scheme, such as the Random Sample Consensus Algorithm<br>( information of the environment to eliminate dynamic features,<br>and a prevalent approach is the maximum consistency<br>scheme, such as the Random Sample Consensus Algorithm<br>(RANSAC [8]). In addition, many visual SALM systems us and a prevalent approach is the maximum consistency<br>scheme, such as the Random Sample Consensus Algorithm<br>(RANSAC [8]). In addition, many visual SALM systems use<br>multi-sensor fusion to detect dynamic targets in the<br>environ scheme, such as the Random Sample Consensus Algorithm (RANSAC [8]). In addition, many visual SALM systems use multi-sensor fusion to detect dynamic targets in the environment, such as ORB-SLAM3 [9]. However, these methods (RANSAC [8]). In addition, many visual SALM systems use<br>multi-sensor fusion to detect dynamic targets in the<br>environment, such as ORB-SLAM3 [9]. However, these<br>methods are effective only when dynamic objects are few,<br>and t multi-sensor fusion to detect dynamic targets in the<br>environment, such as ORB-SLAM3 [9]. However, these<br>methods are effective only when dynamic objects are few,<br>and they fail to capture high-level information about the<br>env environment, such as ORB-SLAM3 [9]. However, these<br>methods are effective only when dynamic objects are few,<br>and they fail to capture high-level information about the<br>environment, resulting in an insufficient understanding exthods are effective only when dynamic objects are few,<br>d they fail to capture high-level information about the<br>vironment, resulting in an insufficient understanding of the<br>rroundings. Dynamic visual SLAM based on deep le and they fail to capture high-level information about the environment, resulting in an insufficient understanding of the surroundings. Dynamic visual SLAM based on deep learning, also known as semantic SLAM, is capable of environment, resulting in an insufficient understanding of the surroundings. Dynamic visual SLAM based on deep learning, also known as semantic SLAM, is capable of acquiring both geometric information about unfamiliar envi surroundings. Dynamic visual SLAM based on deep learning, also known as semantic SLAM, is capable of acquiring both geometric information about unfamiliar environments and the motion states of robots. Moreover, it can dete also known as semantic SLAM, is capable of acquiring both<br>geometric information about unfamiliar environments and<br>the motion states of robots. Moreover, it can detect and<br>recognize targets in the surroundings, allowing for

geometric information about unfamiliar environments and<br>the motion states of robots. Moreover, it can detect and<br>recognize targets in the surroundings, allowing for the<br>filtering out of dynamic feature points (dfp). This c the motion states of robots. Moreover, it can detect and recognize targets in the surroundings, allowing for the filtering out of dynamic feature points (dfp). This capability enables the robot to enhance its comprehension recognize targets in the surroundings, allowing for the filtering out of dynamic feature points (dfp). This capability enables the robot to enhance its comprehension of its surroundings. Moreover it also allows robots to f

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

environmental awareness [2]. So, vision-based simultaneous<br>
localization and map building techniques have been widely<br>
studied and applied to robot navigation [3], unmanned<br>
driving [4] and virtual reality [5].<br>
In practic localization and map building techniques have been widely<br>studied and applied to robot navigation [3], unmanned<br>driving [4] and virtual reality [5].<br>In practical applications, visual SLAM needs to be<br>surroundings. Dy<br>star (KJ1704072).<br>
Zhangfang Hu is a Professor at the Key Laboratory of Optical died and applied to robot navigation [3], unmanned<br>
iving [4] and virtual reality [5].<br>
In practical applications, visual SLAM needs to be<br>
environment, resulting<br>
surroundings. Dynam<br>
also known as semant<br>
Manuscript rece driving [4] and virtual reality [5].<br>
In practical applications, visual SLAM needs to be<br>
surroundings. Dyt<br>
also known as ser<br>
Manuscript received on April 4, 2024; revised on August 27, 2024.<br>
This work was supported in In practical applications, visual SLAM needs to be<br>
surroundings. Dyr<br>
Manuscript received on April 4, 2024; revised on August 27, 2024.<br>
This work was supported in part by the Youth Fund Program of the motion states<br>
Nati Manuscript received on April 4, 2024; revised on August 27<br>This work was supported in part by the Youth Fund Pre<br>National Natural Science Foundation of China (Grant No. 61<br>Chongqing Basic Science and Frontier Technology Re also known as a<br>subsorbing the School of China (Grant No. 61703067), the<br>mongting Basic Science Foundation of China (Grant No. 61703067), the<br>mongting Basic Science and Frontier Technology Research Program of Chongqing Alt Manuscript received on April 4, 2024; revised on August 27, 2024. geometric<br>
This work was supported in part by the Youth Fund Program of the motio<br>
National Natural Science Foundation of China (Grant No. 61703067), the<br>
C Finis work was supported in part +, 2021, ivisus on August 2.1, 2024.<br>
This work was supported in part by the Youth Fund Program of the motion stational Natural Science and Frontier Technology Research Program<br>
Correspondi Frame War was supported in part by the Total That the Science Foundation of China (Grant Chongqing Basic Science and Frontier Technology (Grant No. Cstc2017jcyjAX0212), and the Science Research Program of Chongqing Municip From the School of Optoelectronic Technology Research Program<br>
Dengaing Basic Science and Frontier Technology Research Program<br>
The Science and Frontier Technology Research Program<br>
The School of Consequence and Technology

**Engineering Letters**<br>approaches only remove feature points by relying on the hypothesis, leading to poor<br>results of semantic segmentation or object detection, as performance in complex movin<br>demonstrated by methods such a **Engineering Letters**<br>approaches only remove feature points by relying on the hypothesis, leading to poor<br>results of semantic segmentation or object detection, as performance in complex moving<br>demonstrated by methods such **Engineering Letters**<br>
approaches only remove feature points by relying on the hypothesis, leading to poor<br>
results of semantic segmentation or object detection, as performance in complex move<br>
demonstrated by methods such **Engineering Letters**<br>approaches only remove feature points by relying on the hypothesis, leading to poor<br>results of semantic segmentation or object detection, as<br>performance in complex movid<br>alemonstrated by methods such **Engineering Letters**<br>
approaches only remove feature points by relying on the hypothesis, leading to<br>
results of semantic segmentation or object detection, as<br>
performance in complex<br>
demonstrated by methods such as DS-SL **Engineering Letters**<br>approaches only remove feature points by relying on the hypothesis, leading to po<br>results of semantic segmentation or object detection, as performance in complex mo<br>demonstrated by methods such as DS-**Engineering Letters**<br>
approaches only remove feature points by relying on the hypothesis, leading to poor<br>
results of semantic segmentation or object detection, as performance in complex movid<br>
demonstrated by methods suc **Engineering Letters**<br>
approaches only remove feature points by relying on the hypothesis, leading to po<br>
results of semantic segmentation or object detection, as performance in complex m<br>
demonstrated by methods such as D **Engineering Letters**<br>
approaches only remove feature points by relying on the hypothesis, leading to poor<br>
results of semantic segmentation or object detection, as<br>
performance in complex movin,<br>
demonstrated by methods s **Engineering Letters**<br>
approaches only remove feature points by relying on the hypothesis, leading to poor<br>
results of semantic segmentation or object detection, as<br>
performance in complex movin,<br>
demonstrated by methods s approaches only remove feature points by relying on the hypothesis, leading to presults of semantic segmentation or object detection, as performance in complex n<br>demonstrated by methods such as DS-SLAM [10], enhance the ro approaches only remove feature points by relying on the hypothesis, leading to poor<br>results of semantic segmentation or object detection, as performance in complex movi<br>demonstrated by methods such as DS-SLAM [10], enhance approaches only remove feature points by relying on the hypothesis, leading to p<br>results of semantic segmentation or object detection, as performance in complex m<br>demonstrated by methods such as DS-SLAM [10], enhance the r results of semantic segmentation or object detection, as<br>
performance in complex mo<br>
demonstrated by methods such as DS-SLAM [10], enhance the robustness of vis<br>
DynaSLAM [11], and SaD-SLAM [12]. This approach can<br>
better demonstrated by methods such as DS-SLAM [10], enhance the robustness c<br>DynaSLAM [11], and SaD-SLAM [12]. This approach can<br>better human-machine in<br>lead to two main issues: misidentification of stationary<br>feature points as DynaSLAM [11], and SaD-SLAM [12]. This approach can<br>better human-machine interaction<br>lead to two main issues: misidentification of stationary<br>feature points as moving, which reduces the number of useful<br>object detection an lead to two main issues: misidentification of stationary<br>
feature points as moving, which reduces the number of useful<br>
doject detection and semantifeatures and impairs position estimation capability; and<br>
incomplete remov feature points as moving, which reduces the number of useful object detection and semarate<br>tures and impairs position estimation capability; and remove dynamic regions wi<br>incompromises the precision and robustness of this features and impairs position estimation capability; and remove dynamic region<br>incomplete removal of some dynamic objects, which estimation and the de<br>compromises the precision and robustness of this system. In semantic in complete removal of some dynamic objects, which estimation and the development of impromises the precision and robustness of this system. In semantic information rely solely on sfit<br>dition, these methods use segmentation m compromises the precision and robustness of this system. In<br>
semantic information rely solely<br>
addition, these methods use segmentation models such as<br>
Object detection plays a cru<br>
SegNet [13], Mask R-CNN [14], etc. Altho addition, these methods use segmentation models such as<br>
SegNet [13], Mask R-CNN [14], etc. Although these models<br>
computer vision [17], aimed<br>
have high accuracy, they are more complex in structure and<br>
objects. It involv SegNet [13], Mask R-CNN [14], etc. Although these models<br>
computer vision [17]<br>
have high accuracy, they are more complex in structure and<br>
objects. It involves<br>
take longer to process the data, which fails to meet the cri

have high accuracy, they are more complex in structure and<br>
the longer to process the data, which fails to meet the criteria<br>
for real-time performance. Lastly, the conventional ORB<br>
boxes, and determining the<br>
(Oriented F take longer to process the data, which fails to meet the criteria images or videos, marking the for real-time performance. Lastly, the conventional ORB boxes, and determining their cat (Oriented FAST and Rotated BRIEF) fea for real-time performance. Lastly, the conventional ORB boxes, and determinin<br>(Oriented FAST and Rotated BRIEF) feature extraction of deep learning<br>method [15] employs a fixed threshold that is sensitive to learning-based (Oriented FAST and Rotated BRIEF) feature extraction of deep learning methoc<br>method [15] employs a fixed threshold that is sensitive to learning-based object detection<br>changes in environmental lighting conditions. This ach method [15] employs a fixed threshold that is sensitive to learning-based object detectomages in environmental lighting conditions. This achieving impressive result dependency can lead to challenges such as failures in fea changes in environmental lighting conditions. This achieving impressive rest<br>dependency can lead to challenges such as failures in feature Fast R-CNN [18] and the<br>point extraction and redundancy in local feature points. Re dependency can lead to challenges such as failures in feature Fast R-CNN [18] and the YOLC<br>
point extraction and redundancy in local feature points. Researchers have incorporate<br>
To address the aforementioned challenges in point extraction and redundancy in local feature points. Researchers have incorry To address the aforementioned challenges in SLAM into SLAM systems to systems, we have developed a semantic SLAM system entities on the perf To address the aforementioned challenges in SLAM into SLAM systems to<br>systems, we have developed a semantic SLAM system entities on the performance<br>utilizing the ORB-SLAM3 framework, which ensures both F et al. proposed De systems, we have developed a semantic SLAM system<br>
utilizing the ORB-SLAM3 framework, which ensures both F et al. proposed Detect-SLAM<br>
efficiency and reliability in complex dynamic environments. with deep neural network ( utilizing the ORB-SLAM3 framework, which ensures both F et al. proposed Detect-SLAM<br>efficiency and reliability in complex dynamic environments. with deep neural network (D)<br>Firstly, to tackle the problem of image blurrines efficiency and reliability in complex dynamic environments. with deep neural network<br>
Firstly, to tackle the problem of image blurriness resulting This integration enhances the<br>
from camera shake and rapid motion of dynami Firstly, to tackle the problem of image blurriness resulting This integration enhances the from camera shake and rapid motion of dynamic objects, we tasks effectively and reliably propose the implementation of an image enh from camera shake and rapid motion of dynamic objects, we tasks effectively and relial<br>propose the implementation of an image enhancement environments. Detect-SL/<br>module. This module utilizes the DeblurGANv2 network [16] D propose the implementation of an image enhancement environments. Detect-SLAM<br>module. This module utilizes the DeblurGANv2 network [16] DNN-based detectors, simultar<br>to process blurry images, thereby improving image quality module. This module utilizes the DeblurGANv2 network [16] DNN-based detectors, sinco process blurry images, thereby improving image quality tasks: enhancing SLAM read and facilitating the subsequent modules' operation. imp to process blurry images, thereby improving image quality tasks: enhancing SLAM robustn<br>and facilitating the subsequent modules' operation. improving object detection per<br>Furthermore, to alleviate the effects of dynamic en and facilitating the subsequent modules' operation. improving object detection per Furthermore, to alleviate the effects of dynamic entities on semantic maps. In highly dyn the SLAM system, we incorporate a parallel object Furthermore, to alleviate the effects of dynamic entities on semantic maps. In highly<br>the SLAM system, we incorporate a parallel object detection estimated by the Detect-SL<br>thread within the ORB-SLAM3 framework, utilizing the SLAM system, we incorporate a parallel object detection<br>
the production the order of the DRB-SLAM3<br>
framework, utilizing the true trajectory. However,<br>
epipolar constraints and enhanced YOLOv7 to eliminate the Detect-S thread within the ORB-SLAM3 framework, utilizing the true trajectory. Howev epipolar constraints and enhanced YOLOv7 to eliminate the Detect-SLAM fails to yield dfp. In the object detection network to derive some static in epipolar constraints and enhanced YOLOv7 to eliminate the Detect-SLAM fails to yield des<br>dfp. In the object detection threads, we employ the This is because in static scen<br>lightweight YOLOv7 object detection network to der dfp. In the object detection threads, we employ the This is because in static sightweight YOLOv7 object detection network to derive some static information the semantic details from images. Simultaneously, we utilize this lightweight YOLOv7 object detection network to derive<br>some static information<br>semantic details from images. Simultaneously, we utilize this<br>thread to identify the location of the target within the image. overall system's l semantic details from images. Simultaneously, we utilize this<br>
thread to identify the location of the target within the image.<br>
Meditionally, to addexise the low precision issue of the<br>
Meditionally, to addexise the Semant thread to identify the location of the target within<br>Additionally, to address the low precision is:<br>YOLOv7 network, we integrate the SimAM<br>mechanism into its feature extraction process.<br>enhance the traditional ORB feature dditionally, to address the low precision issue of the Semantic segmentation is a m<br>
DLOv7 network, we integrate the SimAM attention utilized within the domain of co<br>
cchanism into its feature extraction process. Lastly, w YOLOv7 network, we integrate the SimAM attention utilized within the domain of mechanism into its feature extraction process. Lastly, we the classification of each pix enhance the traditional ORB feature extraction method mechanism into its feature extraction process. Lastly, we the classification of e<br>enhance the traditional ORB feature extraction method by categories. Different<br>daaptively adjusting the detection threshold of FAST corners enhance the traditional ORB feature extraction method by<br>adaptively adjusting the detection threshold of FAST corners<br>the images. Semantic segmentation<br>ange. A comparison with conventional ORB feature obtaining more precis adaptively adjusting the detection threshold of FAST corners<br>based on the grayscale values of different regions in the into several regions but also climage. A comparison with conventional ORB feature obtaining more precis

based on the grayscale values of different regions in the into several regions but also c<br>
image. A comparison with conventional ORB feature obtaining more precise image<br>
extraction, which uses a fixed threshold, demonstra image. A comparison with conventional ORB feature obtaining more precise image<br>extraction, which uses a fixed threshold, demonstrates that utilized in multiple domains, in<br>our approach produces a greater number of useful f extraction, which uses a fixed threshold, demonst<br>our approach produces a greater number of usef<br>points with a more uniform distribution. This enh<br>ultimately increases the accuracy of subseque<br>estimation tasks.<br>The followi It is<br>
interest accuracy of subsequent pose It is<br>
enced<br>
It is enced<br>
Interest of this article are structured in the RGI<br>
Elow. In Section II, we present some and<br>
or dynamic environments, summarizing segn<br>
recture. In Se mation tasks.<br>
The following parts of this article are structured in the RGB-D camera configuration and the RGB-D camera configuration and the comeras. Mass<br>
mantic SLAMs for dynamic environments, summarizing segmentation

**Exercise 19 Exercise 2018**<br>
hypothesis, leading to poor localization and mapping<br>
performance in complex moving surroundings. In order to<br>
enhance the robustness of visual SLAM systems and enable<br>
better human-machine int **g Letters**<br>hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researcher g Letters<br>hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researchers **g Letters**<br>hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researcher **g Letters**<br>hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researcher **g Letters**<br>hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researcher **g Letters**<br>hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researcher **Exercise 15 Exercise 15 Exercise 15 Exercise 15 Exercise 15 Experimence in complex moving surroundings. In order to enhance the robustness of visual SLAM systems and enable better human-machine interaction, researchers ha g Letters**<br>hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researcher **Letters**<br> **Exerters**<br> **Exerters**<br> **Exerters**<br> **Exerters**<br> **Exerter**<br> **Exerters**<br> **Exerter**<br> **Exerter** hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researchers have propo hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researchers have propo hypothesis, leading to poor localization and mapping<br>performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researchers have propo

performance in complex moving surroundings. In order to<br>enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researchers have proposed<br>Semantic SLAM. This approach utilizes deep learni enhance the robustness of visual SLAM systems and enable<br>better human-machine interaction, researchers have proposed<br>Semantic SLAM. This approach utilizes deep learning-based<br>object detection and semantic segmentation algo better human-machine interaction, researchers have proposed<br>Semantic SLAM. This approach utilizes deep learning-based<br>object detection and semantic segmentation algorithms to<br>remove dynamic regions within images. Subsequen Semantic SLAM. This approach utilizes deep learning-based<br>object detection and semantic segmentation algorithms to<br>remove dynamic regions within images. Subsequently, pose<br>estimation and the development of maps that incorp object detection and semantic segmentation algorithms to<br>remove dynamic regions within images. Subsequently, pose<br>estimation and the development of maps that incorporate<br>semantic information rely solely on sfp.<br>Object dete nove dynamic regions within images. Subsequently, pose<br>imation and the development of maps that incorporate<br>mantic information rely solely on sfp.<br>Object detection plays a crucial role in the domain of<br>mputer vision [17], estimation and the development of maps that incorporate<br>semantic information rely solely on sfp.<br>Object detection plays a crucial role in the domain of<br>computer vision [17], aimed at localizing and classifying<br>objects. It semantic information rely solely on sfp.<br>
Object detection plays a crucial role in the domain of<br>
computer vision [17], aimed at localizing and classifying<br>
objects. It involves locating desired objects within given<br>
image Object detection plays a crucial role in the domain of computer vision [17], aimed at localizing and classifying objects. It involves locating desired objects within given images or videos, marking their positions with bou

computer vision [17], aimed at localizing and classifying<br>objects. It involves locating desired objects within given<br>images or videos, marking their positions with bounding<br>boxes, and determining their categories. With the objects. It involves locating desired objects within given<br>images or videos, marking their positions with bounding<br>boxes, and determining their categories. With the progression<br>of deep learning methodologies, numerous deep images or videos, marking their positions with bounding<br>boxes, and determining their categories. With the progression<br>of deep learning methodologies, numerous deep<br>learning-based object detection algorithms have emerged,<br>a boxes, and determining their categories. With the progression<br>of deep learning methodologies, numerous deep<br>learning-based object detection algorithms have emerged,<br>achieving impressive results in this field. Examples incl of deep learning methodologies, numerous deep<br>learning-based object detection algorithms have emerged,<br>achieving impressive results in this field. Examples include<br>Fast R-CNN [18] and the YOLO [19] series of algorithms.<br>Re learning-based object detection algorithms have emerged,<br>achieving impressive results in this field. Examples include<br>Fast R-CNN [18] and the YOLO [19] series of algorithms.<br>Researchers have incorporated object detection a achieving impressive results in this field. Examples include<br>Fast R-CNN [18] and the YOLO [19] series of algorithms.<br>Researchers have incorporated object detection algorithms<br>into SLAM systems to mitigate the impact of dyn Fast R-CNN [18] and the YOLO [19] series of algorithms.<br>
Researchers have incorporated object detection algorithms<br>
into SLAM systems to mitigate the impact of dynamic<br>
entities on the performance of these systems. In 2018 Researchers have incorporated object detection algorithms<br>into SLAM systems to mitigate the impact of dynamic<br>entities on the performance of these systems. In 2018, Zhong<br>F et al. proposed Detect-SLAM [20], that integrates into SLAM systems to mitigate the impact of dynamic<br>entities on the performance of these systems. In 2018, Zhong<br>F et al. proposed Detect-SLAM [20], that integrates SLAM<br>with deep neural network (DNN)-based object detector entities on the performance of these systems. In 2018, Zhong F et al. proposed Detect-SLAM [20], that integrates SLAM with deep neural network (DNN)-based object detectors. This integration enhances the ability of robots t F et al. proposed Detect-SLAM [20], that integrates SLAM<br>with deep neural network (DNN)-based object detectors.<br>This integration enhances the ability of robots to perform<br>tasks effectively and reliably in unfamiliar and dy with deep neural network (DNN)-based object detectors.<br>This integration enhances the ability of robots to perform<br>tasks effectively and reliably in unfamiliar and dynamic<br>environments. Detect-SLAM combines SLAM with<br>DNN-ba This integration enhances the ability of robots to perform<br>tasks effectively and reliably in unfamiliar and dynamic<br>environments. Detect-SLAM combines SLAM with<br>DNN-based detectors, simultaneously accomplishing three<br>tasks tasks effectively and reliably in unfamiliar and dynamic<br>environments. Detect-SLAM combines SLAM with<br>DNN-based detectors, simultaneously accomplishing three<br>tasks: enhancing SLAM robustness in dynamic environments,<br>improv vironments. Detect-SLAM combines SLAM with<br>NN-based detectors, simultaneously accomplishing three<br>ks: enhancing SLAM robustness in dynamic environments,<br>proving object detection performance, and constructing<br>mantic maps. I DNN-based detectors, simultaneously accomplishing three tasks: enhancing SLAM robustness in dynamic environments, improving object detection performance, and constructing semantic maps. In highly dynamic scenes, the trajec tasks: enhancing SLAM robustness in dynamic environments,<br>improving object detection performance, and constructing<br>semantic maps. In highly dynamic scenes, the trajectory<br>estimated by the Detect-SLAM system closely approxi improving object detection performance, and constructing<br>semantic maps. In highly dynamic scenes, the trajectory<br>estimated by the Detect-SLAM system closely approximates<br>the true trajectory. However, compared to ORB-SLAM2,

semantic maps. In highly dynamic scenes, the trajectory<br>estimated by the Detect-SLAM system closely approximates<br>the true trajectory. However, compared to ORB-SLAM2,<br>Detect-SLAM fails to yield desirable results in static s estimated by the Detect-SLAM system closely approximates<br>the true trajectory. However, compared to ORB-SLAM2,<br>Detect-SLAM fails to yield desirable results in static scenes.<br>This is because in static scenes, Detect-SLAM fil the true trajectory. However, compared to ORB-SLAM2,<br>
Detect-SLAM fails to yield desirable results in static scenes.<br>
This is because in static scenes, Detect-SLAM filters out<br>
some static information that is beneficial fo Detect-SLAM fails to yield desirable results in static scenes.<br>This is because in static scenes, Detect-SLAM filters out<br>some static information that is beneficial for camera pose<br>estimation and subsequent mapping, thereby This is because in static scenes, Detect-SLAM filters out<br>some static information that is beneficial for camera pose<br>estimation and subsequent mapping, thereby affecting the<br>overall system's localization accuracy.<br>Semantic me static information that is beneficial for camera pose<br>timation and subsequent mapping, thereby affecting the<br>erall system's localization accuracy.<br>Semantic segmentation is a method of image segmentation<br>lized within the estimation and subsequent mapping, thereby affecting the overall system's localization accuracy.<br>
Semantic segmentation is a method of image segmentation utilized within the domain of computer vision. It focuses on the cla overall system's localization accuracy.<br>
Semantic segmentation is a method of image segmentation<br>
utilized within the domain of computer vision. It focuses on<br>
the classification of each pixel in an image into predefined<br> Semantic segmentation is a method of image segmentation<br>utilized within the domain of computer vision. It focuses on<br>the classification of each pixel in an image into predefined<br>categories. Different from traditional image

utilized within the domain of computer vision. It focuses on<br>the classification of each pixel in an image into predefined<br>categories. Different from traditional image segmentation<br>techniques, semantic segmentation not only the classification of each pixel in an image into predefined<br>categories. Different from traditional image segmentation<br>techniques, semantic segmentation not only divides an image<br>into several regions but also classifies ea categories. Different from traditional image segmentation techniques, semantic segmentation not only divides an image into several regions but also classifies each pixel, thereby obtaining more precise image segmentation r techniques, semantic segmentation not only divides an image<br>into several regions but also classifies each pixel, thereby<br>obtaining more precise image segmentation results. It is<br>utilized in multiple domains, including auto into several regions but also classifies each pixel, thereby<br>obtaining more precise image segmentation results. It is<br>utilized in multiple domains, including autonomous driving,<br>medical image analysis, and robotic vision.<br> obtaining more precise image segmentation results. It is<br>utilized in multiple domains, including autonomous driving,<br>medical image analysis, and robotic vision.<br>In 2018, Bescos et al. introduced the DynaSLAM system.<br>It is utilized in multiple domains, including autonomous driving,<br>medical image analysis, and robotic vision.<br>In 2018, Bescos et al. introduced the DynaSLAM system.<br>It is founded on ORB-SLAM2 framework. This system<br>encompasses i medical image analysis, and robotic vision.<br>
In 2018, Bescos et al. introduced the DynaSLAM system.<br>
It is founded on ORB-SLAM2 framework. This system<br>
encompasses interfaces designed for monocular, stereo, and<br>
RGB-D came In 2018, Bescos et al. introduced the DynaSLAM system.<br>It is founded on ORB-SLAM2 framework. This system<br>encompasses interfaces designed for monocular, stereo, and<br>RGB-D camera configurations. When we utilize monocular<br>and It is founded on ORB-SLAM2 framework. This system<br>encompasses interfaces designed for monocular, stereo, and<br>RGB-D camera configurations. When we utilize monocular<br>and stereo cameras, Mask R-CNN is employed to perform<br>segm requirements.

**Engineering Letters**<br>The same year, C. Yu et al. introduced DS-SLAM, a semantic information pertaining<br>thod derived from ORB-SLAM2. Its main innovation lies including labels and positions.<br>the addition of a independent re **Engineering Letters**<br>The same year, C. Yu et al. introduced DS-SLAM, a semantic information pertain<br>method derived from ORB-SLAM2. Its main innovation lies including labels and positions<br>in the addition of a independent r **Engineering Letters**<br>
The same year, C. Yu et al. introduced DS-SLAM, a<br>
method derived from ORB-SLAM2. Its main innovation lies<br>
in the addition of a independent real-time semantic made to the traditional OF<br>
segmentatio **Engineering Letters**<br>The same year, C. Yu et al. introduced DS-SLAM, a<br>mematic informatic<br>method derived from ORB-SLAM2. Its main innovation lies<br>including labels and<br>in the addition of a independent real-time semantic ma **Engineering Letters**<br>The same year, C. Yu et al. introduced DS-SLAM, a semantic information pertaini<br>method derived from ORB-SLAM2. Its main innovation lies including labels and positions<br>in the addition of a independent **Engineering Letters**<br>
The same year, C. Yu et al. introduced DS-SLAM, a<br>
method derived from ORB-SLAM2. Its main innovation lies<br>
in the addition of a independent real-time semantic made to the traditional ORB f<br>
segmenta **Engineering Letters**<br>
The same year, C. Yu et al. introduced DS-SLAM, a semantic information pertaini<br>
method derived from ORB-SLAM2. Its main innovation lies including labels and positions<br>
in the addition of a independe **Engineering L**<br>
The same year, C. Yu et al. introduced DS-SLAM, a<br>
method derived from ORB-SLAM2. Its main innovation lies<br>
in the addition of a independent real-time semantic made<br>
segmentation module within the framewor **Engineering Letters**<br>
The same year, C. Yu et al. introduced DS-SLAM, a<br>
semantic information pertaini<br>
thod derived from ORB-SLAM2. Its main innovation lies<br>
including labels and positions<br>
the addition of a independent **Engineering Letters**<br>
The same year, C. Yu et al. introduced DS-SLAM, a<br>
method derived from ORB-SLAM2. Its main innovation lies<br>
including labels and<br>
in the addition of a independent real-time semantic made to the tradi The same year, C. Yu et al. introduced DS-SLAM, a semantic information pertamethod derived from ORB-SLAM2. Its main innovation lies including labels and position in the addition of a independent real-time semantic made to The same year, C. Yu et al. introduced DS-SLAM, a semantic information p<br>method derived from ORB-SLAM2. Its main innovation lies including labels and poin<br>in the addition of a independent real-time semantic made to the tr

The same year, C. Yu et al. introduced DS-SLAM, a semantic information pertain<br>
method derived from ORB-SLAM2. Its main innovation lies<br>
in the addition of a independent real-time semantic made to the traditional ORB<br>
seg method derived from ORB-SLAM2. Its main innovation lies<br>
including labels and positions.<br>
in the addition of a independent real-time semantic made to the traditional ORB feat<br>
segmentation module within the framework of O in the addition of a independent real-time semantic made to the traditional ORB feat<br>segmentation module within the framework of ORB-SLAM2. by employing adaptive thre<br>environment and creating a dense semantic octree map i segmentation module within the framework of ORB-SLAM2. by employing adaptive thres<br>
This thread is capable of removing dynamic objects from the robustness of feature point extra<br>
econtrimument and creating a dense semanti This thread is capable of removing dynamic objects from the robustness of feature point extra<br>
environment and creating a dense semantic octree map information obtained from the<br>
containing environmental semantic informati environment and creating a dense semantic octree map<br>
containing environmental semantic information, enabling the<br>
conbined with epipolar c<br>
crobot to perform higher-level tasks.<br>
that are linked to dynamic<br>
orbot to perfo containing environmental semantic information, enabling the<br>
robot to perform higher-level tasks.<br>
The above-mentioned methods have improved the we will exclusively emplo<br>
property of SLAM systems to some extent, but mecha The above-mentioned methods have improved the we will exclusively employ state<br>
property of SLAM systems to some extent, but mechanically<br>
ure will exclusively employ state<br>
premoving dynamic objects can result in the loss The above-mentioned methods have improved the we will exclusively employ statt<br>property of SLAM systems to some extent, but mechanically<br>removing dynamic objects can result in the loss of many<br>cabon, Y et al. proposed the property of SLAM systems to some extent, but mechanically<br>
removing dynamic objects can result in the loss of many<br>
usable feature points in the system. To tackle this problem,<br>
Cabon, Y et al. proposed the SLAMANTIC syste removing dynamic objects can result in the loss of many<br>
usable feature points in the system. To tackle this problem,<br>
cabon, Y et al. proposed the SLAMANTIC system [21],<br>
unito does not require motion detection. Instead, usable feature points in the system. To tackle this prob<br>Cabon, Y et al. proposed the SLAMANTIC system [<br>which does not require motion detection. Instead<br>introduces confidence by assigning different probabilitie<br>motion to bon, Y et al. proposed the SLAM aNTIC system [21],<br>
ich does not require motion detection. Instead, it<br>
roduces confidence by assigning different probabilities of<br>
to coach object to ascertain whether the object is in a<br> which does not require motion detection. Instead, it<br>introduces confidence by assigning different probabilities of<br>motion to each object to ascentian whether the object is in a<br>ability to distinguish between objects that introduces confidence by assigning different probabilities of<br>
motion to each object to ascertain whether the object is in a<br>
state of motion. As a result, this methodology possesses the<br>
ast dynamic, while they are, in re motion to each object to ascertain whether the object is in a<br>
state of motion. As a result, this methodology possesses the<br>
ability to distinguish between objects that might be regarded<br>
ability to distinguish between obj state of motion. As a result, this methodology possesses the<br>ability to distinguish between objects that might be regarded<br>as dynamic, while they are, in reality, stationary. In addition,<br>This system integrates semantic la

Network ability to distinguish between objects that might be regarded<br>
as dynamic, while they are, in reality, stationary. In addition,<br>
This system integrates semantic label distribution with the<br>
consistency of map point as dynamic, while they are, in reality, stationary. In addition,<br>
This system integrates semantic label distribution with the<br>
consistency of map point observations to vealuate the<br>
reliability of each 3D measurement point This system integrates semantic label distribution with<br>consistency of map point observations to evaluate<br>reliability of each 3D measurement point. Subsequently,<br>information is utilized for pose estimation and subsequ<br>map Unifized for pose estimation and subsequent<br>
in steps.<br>
semantic SLAM outperforms traditional<br>
in overall performance. However, some<br>
solely on improving accuracy while<br>
In the<br>
I-time capabilities, while others exhibit go This section offers a comprehensive analysis of the system<br>
understand thos cours solely on improving accuracy while<br>
understand SLAM in overall performance. However, some<br>
erlooking real-time epapolities, while others exh In summary, semantic SLAM outperforms traditional<br>
in overall performance. However, some<br>
methods focus solely on improving accuracy while<br>
overlooking real-time capabilities, while others exhibit good<br>
real-time performa VISUAL SLAM in overall performance. However, some<br>
methods focus solely on improving accuracy while<br>
overlooking real-time epapelitities, while other exhibit good<br>
real-time performance but lower accuracy and robustness.<br>

methods focus solely on improving accuracy while<br>overlooking real-time epablities, while others exhibit good<br>real-time performance but lower accuracy and robustness.<br>Therefore, enhancing the accuracy of SLAM systems while<br> overlooking real-time capabilities, while others exhibit good<br>
real-time performance but lower accuracy and robustness.<br>
Therefore, enhancing the accuracy of SLAM systems while<br>
maintaining a certain level of real-time pro real-time performance but lower accuracy and robustness.<br>
Therefore, enhancing the accuracy of SLAM systems while<br>
maintaining a certain level of real-time property constitutes a<br>
significant area of research.<br>
III. SYSTEM Therefore, enhancing the accuracy of SLAM systems while<br>
maintaining a certain level of real-time property constitutes a<br>
in environments characterize<br>
in environments characterize<br>
in environments characterize<br>
in enviro maintaining a certain level of real-time property constitutes a<br>
in environments characte<br>
in environments characte<br>
in environments characte<br>
in environments characte<br>
inforculations are improvements of the system propose significant area of research.<br>
III. SYSTEM INTRODUCTION<br>
III. SYST III. SYSTEM INTRODUCTION<br>
introduces an image<br>
This section offers a comprehensive analysis of the system<br>
put forward in the present study. First, we present the<br>
put forward in the present study. First, we present the<br>
i This section offers a comprehensive analysis of the system images obtained from the camer<br>put forward in the present study. First, we present the The DeblurGANv2 network<br>improvements of the system proposed within this pap This section offers a comprehensive analysis of the system<br>put forward in the present study. First, we present the<br>improvements of the system proposed within this paper on<br>DeblurGAI<br>the ORB-SLAM3 framework, including the *A. Framework of our system*<br> *A. Framework, including the incorporation of*<br> *A. Framework, including the incorporation of*<br> *A. Framework, including the incorporation of*<br> *A. Examework of implementing the image enhancem* EXECUTE: A strained in Fig. 1, ORB-SLAM3 is an open-source and foother and the substrated in Fig. 1, ORB-SLAM3 is an open-source of the proposed in the substrated in Fig. 1, ORB-SLAM3 is an open-source pre-transition of s and image enhancement module, the introduction of adaptive<br>
interded backbone network, resulting<br>
thresholding, and the enhanced YOLOv7 network. Then, the<br>
interded variables of implementing the image enhancement module<br>
a mapping thread, and a loop closing thread responsible for<br>methods of implementing the enhanced YOLOv7 network. Then, the<br>discompared to DeblurGAN. I<br>methods of implementing the image enhancement module<br>of performance, meet

methols of implementing the image enhancement module<br>
methods of implementing the image enhancement module<br>
on the traditional<br>
on the traditional<br>
ORB feature extraction are explained in detail. Lastly, the<br>
DeblurGANv2 n independent and using adaptive thresholding to improve the traditional<br>
or a discriminator. The network and<br>
ORB feature extraction are explained in detail. Lastly, the<br>
ORB feature extraction are explained in detail. Last ORB feature extraction are explained in detail. Lastly, the systems. The interwork actionues<br>paper elaborates on the enhancement of the YOLOv7 a discriminator. The generator entwork and its integration with polar constrai paper elaborates on the enhancement of the YOLOv7<br>
paper elaborates on the enhancement of the YOLOv7<br>
adiscriminator. The generator enterwork and its integration with polar constraints to propose a<br>
Metwork (FPN) structur Framework and its integration with polar constraints to propose a<br>
method for filtering dynamic feature points.<br>
The branches and fuses the quality of produced in Fig. 1, ORB-SLAM3 is an open-source<br>
Method is integration Framework of our system<br>
A. Framework of our system<br>
A. Framework of our system<br>
As illustrated in Fig. 1, ORB-SLAM3 is an open-source<br>
The quality of produced im<br>
relativistic discriminator<br>
NS system characterized by a t A. Framework of our system<br>
As illustrated in Fig. 1, ORB-SLAM3 is an open-source<br>
the quality of produced ima<sub>n</sub><br>
relativistic discriminator<br>
NS system characterized by a tracking thread, a local<br>
mapping thread, and a lo *A. Framework of our system*<br>
relativistic discriminator<br>
As illustrated in Fig. 1, ORB-SLAM3 is an open-source<br>
fluendion is implemented. A<br>
VS system characterized by a tracking thread, a local<br>
mapping thread, and a lo

g Letters<br>
semantic information pertaining to entities in images,<br>
including labels and positions. Improvements have been<br>
made to the traditional ORB feature point extraction method<br>
by employing adaptive thresholding, en g Letters<br>
semantic information pertaining to entities in images,<br>
including labels and positions. Improvements have been<br>
made to the traditional ORB feature point extraction method<br>
by employing adaptive thresholding, en g Letters<br>semantic information pertaining to entities in images,<br>including labels and positions. Improvements have been<br>made to the traditional ORB feature point extraction method<br>by employing adaptive thresholding, enhanc **g Letters**<br>semantic information pertaining to entities in images,<br>including labels and positions. Improvements have been<br>made to the traditional ORB feature point extraction method<br>by employing adaptive thresholding, enha g Letters<br>semantic information pertaining to entities in images,<br>including labels and positions. Improvements have been<br>made to the traditional ORB feature point extraction method<br>by employing adaptive thresholding, enhanc g Letters<br>semantic information pertaining to entities in images,<br>including labels and positions. Improvements have been<br>made to the traditional ORB feature point extraction method<br>by employing adaptive thresholding, enhanc **g Letters**<br>
semantic information pertaining to entities in images,<br>
including labels and positions. Improvements have been<br>
made to the traditional ORB feature point extraction method<br>
by employing adaptive thresholding, **Example 18 Example 18 Concerned to the subsequent tasks,**<br> **Example 18 Concerned to the traditional ORB feature point extraction method**<br>
by employing adaptive thresholding, enhancing the<br>
robustness of feature point extr **g Letters**<br>
semantic information pertaining to entities in images,<br>
including labels and positions. Improvements have been<br>
made to the traditional ORB feature point extraction method<br>
by employing adaptive thresholding,



Faster Lawrence Cause of SLAM and the scene can also cause blurriness.<br>
Fast-moving objects in the scene can also cause blurriness.<br>
The scene can also cause blurriness.<br>
Many previous SLAM methods have overlooked these is Full BA<br>
Full BA<br>
Fig.1. Framework of ORB-SLAM3<br>
Fig.1. Framework of ORB-SLAM3<br>
B. *Image Enhancement Module*<br>
In the course of a mobile robot's movement, camera shake<br>
inevitably occurs, leading to blurry images. Additio Full BA<br>
Fig.1. Framework of ORB-SLAM3<br>
Fig.1. Framework of ORB-SLAM3<br>
B. Image Enhancement Module<br>
In the course of a mobile robot's movement, camera shake<br>
inevitably occurs, leading to blurry images. Additionally,<br>
fast Full BA<br>
Fig.1. Framework of ORB-SLAM3<br>
Fig.1. Framework of ORB-SLAM3<br>
B. Image Enhancement Module<br>
In the course of a mobile robot's movement, camera shake<br>
inevitably occurs, leading to blurry images. Additionally,<br>
fast Full BA<br>
Fig.1. Framework of ORB-SLAM3<br>
Fig.1. Framework of ORB-SLAM3<br>
B. Image Enhancement Module<br>
In the course of a mobile robot's movement, camera shake<br>
inevitably occurs, leading to blurry images. Additionally,<br>
fast Fig.1. Framework of ORB-SLAM3.<br> **Example 18.** Image Enhancement Module<br>
In the course of a mobile robot's movement, camera shake<br>
inevitably occurs, leading to blurry images. Additionally,<br>
fast-moving objects in the scene Fig.1. Framework of ORB-SLAM3<br> *B. Image Enhancement Module*<br>
In the course of a mobile robot's movement, camera shake<br>
inevitably occurs, leading to blurry images. Additionally,<br>
fast-moving objects in the scene can also network. 3. *Image Enhancement Module*<br>In the course of a mobile robot's movement, camera shake<br>evitably occurs, leading to blurry images. Additionally,<br>st-moving objects in the scene can also cause blurriness.<br>any previous SLAM me *B. Image Enhancement Module*<br>In the course of a mobile robot's movement, camera shake<br>inevitably occurs, leading to blurry images. Additionally,<br>fast-moving objects in the scene can also cause blurriness.<br>Many previous S In the course of a mobile robot's movement, camera shake<br>inevitably occurs, leading to blurry images. Additionally,<br>fast-moving objects in the scene can also cause blurriness.<br>Many previous SLAM methods have overlooked the inevitably occurs, leading to blurry images. Additionally,<br>fast-moving objects in the scene can also cause blurriness.<br>Many previous SLAM methods have overlooked these issues,<br>resulting in poor robustness and accuracy of S

fast-moving objects in the scene can also cause blurriness.<br>Many previous SLAM methods have overlooked these issues,<br>resulting in poor robustness and accuracy of SLAM systems<br>in environments characterized by high dynamics. Many previous SLAM methods have overlooked these issues,<br>resulting in poor robustness and accuracy of SLAM systems<br>in environments characterized by high dynamics. This paper<br>introduces an image enhancement module into the<br> resulting in poor robustness and accuracy of SLAM systems<br>in environments characterized by high dynamics. This paper<br>introduces an image enhancement module into the<br>framework of ORB-SLAM3. This module preprocesses<br>images o in environments characterized by high dynamics. This paper<br>introduces an image enhancement module into the<br>framework of ORB-SLAM3. This module preprocesses<br>images obtained from the camera using the DeblurGANv2<br>network.<br>The introduces an image enhancement module into the<br>framework of ORB-SLAM3. This module preprocesses<br>images obtained from the camera using the DeblurGANv2<br>network.<br>The DeblurGANv2 network is an improvement over<br>DeblurGANv2 uti framework of ORB-SLAM3. This module preprocesses<br>images obtained from the camera using the DeblurGANv2<br>network.<br>The DeblurGANv2 network is an improvement over<br>DeblurGANv2 utilizes the lightweight MobileNet [23] as its<br>back images obtained from the camera using the DeblurGANv2<br>network.<br>The DeblurGANv2 network is an improvement over<br>DeblurGAN [22], achieving better results. Furthermore,<br>DeblurGANv2 utilizes the lightweight MobileNet [23] as it network.<br>The DeblurGANv2 network is an improvement over<br>DeblurGAN [22], achieving better results. Furthermore,<br>DeblurGANv2 utilizes the lightweight MobileNet [23] as its<br>backbone network, resulting in a 20x speed improveme The DeblurGANv2 network is an improvement over<br>DeblurGAN [22], achieving better results. Furthermore,<br>DeblurGANv2 utilizes the lightweight MobileNet [23] as its<br>backbone network, resulting in a 20x speed improvement<br>compar DeblurGAN [22], achieving better results. Furthermore,<br>DeblurGANv2 utilizes the lightweight MobileNet [23] as its<br>backbone network, resulting in a 20x speed improvement<br>compared to DeblurGAN. It exhibits good real-time<br>per DeblurGANv2 utilizes the lightweight MobileNet [23] as its backbone network, resulting in a 20x speed improvement compared to DeblurGAN. It exhibits good real-time performance, meeting the real-time requirements of SLAM sy backbone network, resulting in a 20x speed improvement<br>compared to DeblurGAN. It exhibits good real-time<br>performance, meeting the real-time requirements of SLAM<br>systems. The network architecture is illustrated in Fig. 3. T performance, meeting the real-time requirements of SLAM<br>systems. The network architecture is illustrated in Fig. 3. The<br>DeblurGANv2 network consists primarily of a generator and<br>a discriminator. The generator employs the F systems. The network architecture is illustrated in Fig. 3. The<br>DeblurGANv2 network consists primarily of a generator and<br>a discriminator. The generator employs the Feature Pyramid<br>Network (FPN) structure, which gathers fe DeblurGANv2 network consists primarily of a generator and<br>a discriminator. The generator employs the Feature Pyramid<br>Network (FPN) structure, which gathers feature outputs from<br>five branches and fuses them through upsampli a discriminator. The generator employs the Feature Pyramid<br>Network (FPN) structure, which gathers feature outputs from<br>five branches and fuses them through upsampling to improve<br>the quality of produced images. In the discr

Network (FPN) structure, which gathers feature outputs from<br>five branches and fuses them through upsampling to improve<br>the quality of produced images. In the discriminator part, a<br>relativistic discriminator utilizing a lea five branches and fuses them through upsampling to improve<br>the quality of produced images. In the discriminator part, a<br>relativistic discriminator utilizing a least-squares loss<br>function is implemented. Additionally, it in the quality of produced images. In the discriminator part, a relativistic discriminator utilizing a least-squares loss function is implemented. Additionally, it integrates global and local scale discriminator losses, ensur

**Engineering Letters**<br>blurring induced by camera shake and rapid object TUM dataset. The outcomes,<br>movements. Subsequently, the deblurred frames are fed into that the image enhancemen<br>feature extraction and object detectio **Engineering Letters**<br>blurring induced by camera shake and rapid object TUM dataset. The outcomes, il<br>movements. Subsequently, the deblurred frames are fed into that the image enhancement if<br>feature extraction and object d **Engineering Letters**<br>blurring induced by camera shake and rapid object TUM dataset. The outcomes<br>movements. Subsequently, the deblurred frames are fed into that the image enhancemer<br>feature extraction and object detectio Engineering Letters<br>
blurring induced by camera shake and rapid object TUM dataset. The outcomes, ill<br>
movements. Subsequently, the deblurred frames are fed into that the image enhancement m<br>
feature extraction and object

TUM dataset. The outcomes, illustrated in Fig.4, substantiate<br>that the image enhancement module utilized in this study<br>significantly ameliorates issues of image blurring attributed<br>to camera shake and object motion. **TUM dataset.** The outcomes, illustrated in Fig.4, substantiate that the image enhancement module utilized in this study significantly ameliorates issues of image blurring attributed to camera shake and object motion. **Solution:**<br>TUM dataset. The outcomes, illustrated in Fig.4, substantiate<br>that the image enhancement module utilized in this study<br>significantly ameliorates issues of image blurring attributed<br>to camera shake and object m **TUM dataset.** The outcomes, illustrated in Fig.4, substantiate<br>that the image enhancement module utilized in this study<br>significantly ameliorates issues of image blurring attributed<br>to camera shake and object motion.





**Engineering Letters**<br>*C. ORB feature extraction based on adaptive thresholding* First, establish a thr<br>In the ORB-SLAM3 system, ORB (Oriented FAST and grayscale values in the<br>otated BRIEF) feature points are utilized, whi **Engineering Letters**<br>
C. *ORB feature extraction based on adaptive thresholding*<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and<br>
tated BRIEF) feature points are utilized, which consist of<br>
iented FAST corners and BRIEF **Engineering Letters**<br>
C. *ORB feature extraction based on adaptive thresholding*<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and<br>
Rotated BRIEF) feature points are utilized, which consist of<br>
oriented FAST comers are ut **Engineering Letters**<br>
C. *ORB feature extraction based on adaptive thresholding*<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and<br>
Rotated BRIEF) feature points are utilized, which consist of<br>
Oriented FAST corners and B **Engineering Letters**<br>
C. *ORB feature extraction based on adaptive thresholding* First, establish a threshold v<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and anal prayscale values in the range [<br>
Rotated BRIEF) featur **Engineering Letters**<br>
C. *ORB feature extraction based on adaptive thresholding* First, establish a threshold In the ORB-SLAM3 system, ORB (Oriented FAST and and [t+1, L-1]. Let *S<sub>l</sub>* and *S*<br>
Rotated BRIEF) feature poi **Engineering Letters**<br>
C. *ORB feature extraction based on adaptive thresholding*<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and grayscale values in the range [0, L<br>
Rotated BRIEF) feature points are utilized, which con **Engineering Letters**<br> *C. ORB feature extraction based on adaptive thresholding*<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and grayscale values in the range [<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and and [*t*+ **Engineering Letters**<br> *C. ORB feature extraction based on adaptive thresholding* First, establish a threshol<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and and [*t*+*I*, *L*-1]. Let *S*<sub>*i*</sub> and<br>
Rotated BRIEF) feature *C. ORB feature extraction based on adaptive thresholding* First, establish<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and and  $[t+1, L-1]$ . I<br>
Rotated BRIEF) feature points are utilized, which consist of probability dist *C. ORB feature extraction based on adaptive thresholding* First, establish a threshold value In the ORB-SLAM3 system, ORB (Oriented FAST and and  $[t+l, L-l]$ . Let  $S_l$  and  $S_2$  re Oriented FAST corners and BRIEF descriptors *C. ORB feature extraction based on adaptive thresholding* First, establish a threshold of In the ORB-SLAM3 system, ORB (Oriented FAST and  $[f+l, L-l]$ . Let  $S_l$  and  $S_l$  oriented FAST corners are utilized, which consist of p C. *ORB feature extraction based on adaptive thresholding*<br>
In the ORB-SLAM3 system, ORB (Oriented FAST and<br>
Rotated BRIEF) feature points are utilized, which consist of<br>
oriented FAST corners and BRIEF descriptors. FAST<br> In the ORB-SLAM3 system, ORB (Oriented FAST and and  $[t+1, L-1]$ . Let S<sub>i</sub> and S<br>
Rotated BRIEF) feature points are utilized, which consist of probability distributions. Concernes primarily detect areas with significant loc Rotated BRIEF) feature points are utilized, which consist of<br>
Oriented FAST corners and BRIEF descriptors. FAST<br>
intensity destributions<br>
corners primarily detect areas with significant local pixel<br>
intensity changes, as Oriented FAST corners and BRIEF descriptors. FAST<br>
corners primarily detect areas with significant local pixel<br>
intensity changes, as illustrated in Fig. 5. The feature point<br>
extraction process, as described in [15], inv corners primarily detect areas with significant local pixel<br>intensity changes, as illustrated in Fig. 5. The feature point<br>extraction process, as described in [15], involves selecting a<br>pixel P in the image with pixel val intensity changes, as illustrated in Fig. 5. The feature point<br>
extraction process, as described in [15], involves selecting a<br>
pixel P in the image with pixel value  $I_p$ . The threshold value<br>
T is set (e.g., 20% of  $I_p$ ) extraction process, as described in [15], involves selecting a<br>
pixel P in the image with pixel value  $I_p$ . The threshold value<br>
T is set (e.g., 20% of  $I_p$ ). Then, 16 pixels are selected around<br>
P, utilizing a radius of pixel P in the image with pixel value  $I_p$ . The threshold value<br>
T is set (e.g., 20% of  $I_p$ ). Then, 16 pixels are selected around<br>
P, utilizing a radius of 3 pixels. If there are N consecutive<br>
points on the circular pat T is set (e.g., 20% of  $I_p$ ). Then, 16 pixels are selected around<br>
P, utilizing a radius of 3 pixels. If there are N consecutive<br>
prece  $P_i$  presents the probably<br>
pixels on the circular path with values exceeding ( $I_p + T$ *P*, utilizing a radius of 3 pixels. If there are *N* consecutive occurrence. Then, let  $P_t = P_0 +$ <br>pixels on the circular path with values exceeding  $(I_p + T)$  or and  $S_2$  can be articulated as:<br>fill is customary to set the pixels on the circular path with values exceeding  $(I_p + T)$  or and  $S_2$  can be articulated a<br>falling blow  $(I_p - T)$ , then P can be classified as a feature point.<br>It is customary to set the value of N to 12, a configuration falling blow ( $I_p$ -T), then P can be classified as a feature point.<br>
It is customary to set the value of N to 12, a configuration<br>
It is customary to set the value of N to 12, a configuration<br>
are 9 and 11, referred to as It is customary to set the value of N to 12, a configuration  $H(S_1)$  =<br>referred to as FAST-12. Other commonly used values for N<br>are 9 and 11, referred to as FAST-9 and FAST-11,<br>respectively). Since FAST corners use a fixe referred to as FAST-12. Other commonly used values for N<br>
are 9 and 11, referred to as FAST-9 and FAST-11,<br>
respectively). Since FAST corners use a fixed threshold<br>
during extraction, only the points with the most signifi are 9 and 11, referred to as FAST-9 and FAST-11,<br>respectively). Since FAST corners use a fixed threshold<br>during extraction, only the points with the most significant<br>This results in the image are selected as corners. Cons pectively). Since FAST corners use a fixed threshold<br> *H*( $S_2$ ) = ring extraction, only the points with the most significant<br>
in sysscale differences in the image are selected as corners.<br>
Consequently, the o<br>
ints withi during extraction, only the points with the most significant<br>
grayscale differences in the image are selected as corners.<br>
This results in the imability to extract other useful feature<br>
points within the image. Furthermor grayscale differences in the image are selected as corners. Consequently, the over<br>This results in the imability to extract other useful feature expressed as the cumulat<br>points within the image. Furthermore, during subseq This results in the inability to extract other useful feature<br>
expressed as the cumulat<br>
points within the image. Furthermore, during subsequent<br>
denoted as:<br>
IF(S) =<br>
preservirong the maximum Harris response value result

points within the image. Furthermore, during subsequent<br>feature extraction using quadtree partitioning, the method of<br>preserving the maximum Harris response value results in all<br>and the method of<br>comers being concentrated feature extraction using quadtree partitioning, the method of<br>preserving the maximum Harris response value results in all<br>corners being concentrated in regions with richer textures. the grayscal corners being concentrated preserving the maximum Harris response value results in all<br>
Recordent and grayscale histogram, ca<br>
This leads to redundant local feature points. If these feature<br>
the two classes. Then, the<br>
points gather on dynamic enti corners being concentrated in regions with richer textures. the grayscale histogram, This leads to redundant local feature points. If these feature the two classes. Then, the points may result in fewer available feature p This leads to redundant local feature points. If these feature<br>
the two classes. Then, the grays<br>
points gather on dynamic entities, removing dynamic feature<br>
the maximum value is denoted as<br>
consecting the solication fai points gather on dynamic entities, removing dynamic feature<br>points may result in fewer available feature points. In severe<br>corresponding to the minimu<br>cases, this can cause localization failure in the SLAM system. Therefo points may result in fewer available feature points. In severce corresponding to the minimum<br>cases, this can cause localization failure in the SLAM system. Therefore, the global optimal the<br>After the above analysis, selec cases, this can cause localization failure in the SLAM system. Therefore, the global optical of the above analysis, selecting ORB features. In order to the robustness of feature point extraction in complex<br>
embance the ro After the above analysis, selecting the appropriate as:<br>
threshold is essential for extracting ORB features. In order to<br>
enhance the robustness of feature point extraction in complex<br>
environments for the SLAM system, we threshold is essential for extracting ORB features. In order to  $T_g = h$ <br>
enhance the robustness of feature point extraction in complex<br>
environments for the SLAM system, we adopt an adaptive<br>
thresholds selection method th enhance the robustness of feature point extraction in complex<br>environments for the SLAM system, we adopt an adaptive<br>thresholds relaction method that is grounded in KSW entropy<br>threshold  $Tg$  can be adaptiv<br>urality value environments for the SLAM system, we adopt an adaptive<br>thresholds selection method that is grounded in KSW entropy<br>threshold  $T_g$  can be adaptive<br>grayscale distribution of the image. The KSW entropy<br>distribution of gravit thresholds selection method that is grounded in KSW entropy<br>value [24], determining the global threshold  $T_g$  based on the distribution of grayscale values grayscale values. The KSW entropy all regions of the image are<br>en value [24], determining the global threshold  $T_g$  based on the distribution of grayscale values is<br>grayscale distribution of the image. The KSW entropy all regions of the image are<br>method refers to calculating the entropy grayscale distribution of the image. The KSW entropy all regions of the image are<br>method refers to calculating the entropy of the grayscale conditions. When local regions<br>histogram of an image and utilizing conditional pr method refers to calculating the entropy of the grayscale<br>
bistogram of an image and utilizing conditional probability to<br>
doscarbistication of grayscale values for objects and<br>
backgrounds in the image, thereby defining histogram of an image and utilizing conditional probability to<br>
describe the distribution of grayscale values for objects and<br>
inadequate, leading to a decre<br>
backgrounds in the image, thereby defining the entropy for<br>
fe describe the distribution of grayscale values for objects and<br>backgrounds in the image, thereby defining the entropy for<br>the constrained both the objects and the backgrounds. The method employed<br>in this paper approximates backgrounds in the image, thereby defining the entropy for<br>both the objects and the backgrounds. The method employed<br>in this paper approximates the probability of each grayscale<br>in this paper are proposition of grayscale both the objects and the backgrounds. The method employed<br>in this paper approximates the probability of each grayscale<br>value to represent the likelihood distribution of grayscale<br>values and extracts global grayscale inform



**Example 18 Example 10**<br>First, establish a threshold value t to divide the image with<br>ayscale values in the range  $[0, L-1]$  into two categories:  $[0, t]$ <br>d  $[t+1, L-1]$ . Let  $S_1$  and  $S_2$  represent the respective pixel<br>bob **g Letters**<br>First, establish a threshold value t to divide the image with grayscale values in the range  $[0, L-1]$  into two categories:  $[0, t]$  and  $[t+1, L-1]$ . Let  $S_1$  and  $S_2$  represent the respective pixel probability **EXECUTE:**<br>First, establish a threshold value t to divide the image with grayscale values in the range [0, L-1] into two categories: [0, t] and [t+1, L-1]. Let  $S_l$  and  $S_2$  represent the respective pixel probability dis **g Letters**<br>
First, establish a threshold value t to divide the image with<br>
grayscale values in the range [0, L-1] into two categories: [0, t]<br>
and [t+1, L-1]. Let  $S_1$  and  $S_2$  represent the respective pixel<br>
probabili **EXECUTE:**<br>
First, establish a threshold value t to divide the image with<br>
grayscale values in the range [0, L-1] into two categories: [0, t]<br>
and [t+1, L-1]. Let  $S_l$  and  $S_2$  represent the respective pixel<br>
probability **1.** threshold value t to divide the image with<br>
the range  $[0, L-I]$  into two categories:  $[0, t]$ <br>
t  $S_I$  and  $S_2$  represent the respective pixel<br>
utions. Consequently, S1 and S2 can be<br>
ollowing formulas:<br>  $S_1 = \{P_0, P_1$ *Sh* a threshold value t to divide the image with<br> *Ss* in the range [0, L-1] into two categories: [0, t]<br> *L* Let  $S_l$  and  $S_2$  represent the respective pixel<br>
tributions. Consequently, S1 and S2 can be<br> *S*<sub>1</sub> = { $P_0,$ **Letters**<br>
First, establish a threshold value t to divide the image with<br>
ayscale values in the range [0, L-1] into two categories: [0, t]<br>  $d$  [t+1, L-1]. Let  $S_1$  and  $S_2$  represent the respective pixel<br>
bobability di First, establish a threshold value t to divide the image with<br>grayscale values in the range [0, L-I] into two categories: [0, t]<br>and [t+1, L-I]. Let  $S_t$  and  $S_2$  represent the respective pixel<br>probability distributions. First, establish a threshold value t to divide the image with<br>grayscale values in the range [0, L-1] into two categories: [0, t]<br>and [t+1, L-1]. Let  $S_t$  and  $S_2$  represent the respective pixel<br>probability distributions. threshold value t to divide the image with<br>
the range [0, L-1] into two categories: [0, t]<br>
S<sub>1</sub> and S<sub>2</sub> represent the respective pixel<br>
oins. Consequently, S1 and S2 can be<br>
lowing formulas:<br>  $\left[ P_0, P_1, P_2, \ldots P_t \right]$  ( *i* to divide the image with<br> *i*] into two categories: [0, t]<br>
esent the respective pixel<br>
ently, S1 and S2 can be<br>
as:<br>  $\ldots P_t$  (1)<br>  $\ldots P_{L-1}$  (2)<br>
(2)<br>
ty of each grayscale level<br>  $+\cdots+P_t$ , the entropy of S<sub>1</sub><br>  $\frac{P$ a threshold value t to divide the image with<br>
the range [0, L-1] into two categories: [0, t]<br>
et S<sub>I</sub> and S<sub>2</sub> represent the respective pixel<br>
butions. Consequently, S1 and S2 can be<br>
following formulas:<br>  $S_1 = \{P_0, P_1, P$ alue t to divide the image with<br> *P*, *L*-*I*] into two categories: [0, *t*]<br>
represent the respective pixel<br>
sequently, S1 and S2 can be<br>
mulas:<br>  $P_1, P_2, \ldots P_l$ } (1)<br>  $P_2, \ldots P_{l-1}$ } (2)<br>
ability of each grayscale leve old value t to divide the image with<br>
gge [0, L-1] into two categories: [0, t]<br>
d S<sub>2</sub> represent the respective pixel<br>
Consequently, S1 and S2 can be<br>
g formulas:<br>  $P_0, P_1, P_2, \dots P_t$ } (1)<br>  $P_{t+2}, P_{t+3}, \dots P_{L-1}$ } (2)<br>
pro 1 a threshold value t to divide the image with<br>
in the range [*0*, *L*-*I*] into two categories: [*0*, *t*]<br>
Let *S<sub>i</sub>* and *S<sub>2</sub>* represent the respective pixel<br>
ibutions. Consequently, S1 and S2 can be<br> *f* ollowing for ralue t to divide the image with<br>  $0, L-l$ ] into two categories: [ $0, t$ ]<br>
represent the respective pixel<br>
sequently, S1 and S2 can be<br>
rmulas:<br>  $2, P_1, P_2, \ldots P_t$ }
(1)<br>  $2, P_{t+3}, \ldots P_{L-1}$ }
(2)<br>
ability of each grayscale lev *i* that a threshold value t to divide the image with<br> *is* in the range [0, L-1] into two categories: [0, t]<br> *J*. Let *S<sub>1</sub>* and *S<sub>2</sub>* represent the respective pixel<br>
stributions. Consequently, S1 and S2 can be<br>
the fo *P* at to divide the image with<br> *P* j into two categories: [0, *t*]<br>
resent the respective pixel<br>
luently, S1 and S2 can be<br>
las:<br>  $\begin{cases}\n\cdot, \dots P_t\end{cases}$  (1)<br>  $\begin{cases}\n\cdot, \dots P_{t-1}\end{cases}$  (2)<br>
ity of each grayscale level<br>  $P_1 +$ reshold value t to divide the image with<br>
r ange [0, L-1] into two categories: [0, t]<br> *t* and S<sub>2</sub> represent the respective pixel<br>
ms. Consequently, S1 and S2 can be<br>
wing formulas:<br>
= { $P_0, P_1, P_2, ..., P_t$ } (1)<br>
When proba bld value t to divide the image with<br>
ge [0, L-1] into two categories: [0, t]<br>  $S_2$  represent the respective pixel<br>
Consequently, S1 and S2 can be<br>
g formulas:<br>  $P_0, P_1, P_2, \dots P_r$ }
(1)<br>  $P_{t+2}, P_{t+3}, \dots P_{L-1}$ }
(2)<br>
proba

$$
S_1 = \{P_0, P_1, P_2, \dots P_t\} \tag{1}
$$

$$
S_2 = \{P_{t+1}, P_{t+2}, P_{t+3}, \dots P_{L-1}\}
$$
 (2)

 $S_1 = \{P_{t+1}, P_{t+2}, P_{t+3}, \dots P_{L-1}\}$  (1)<br>  $S_2 = \{P_{t+1}, P_{t+2}, P_{t+3}, \dots P_{L-1}\}$  (2)<br>
Where  $P_i$  presents the probability of each grayscale level<br>
currence. Then, let  $P_t = P_0 + P_1 + \dots + P_t$ , the entropy of  $S_i$ <br>  $S_2$  can be arti  $S_2 = \{P_{t+1}, P_{t+2}, P_{t+3}, ... P_{L-1}\}$  (2)<br>
Where  $P_i$  presents the probability of each grayscale level<br>
occurrence. Then, let  $P_t = P_0 + P_1 + ... + P_t$ , the entropy of  $S_i$ <br>
and  $S_2$  can be articulated as:<br>  $H(S_1) = -\sum_{i=0}^{t} \frac{P_i}{$ Where  $P_i$  presents the probability of each g<br>occurrence. Then, let  $P_t = P_0 + P_1 + \cdots + P_t$ , the<br>and  $S_2$  can be articulated as:<br> $H(S_1) = -\sum_{i=0}^t \frac{P_i}{P_i} \ln \frac{P_i}{P_t}$ <br> $H(S_2) = -\sum_{i=t+1}^{t-1} \frac{P_i}{1-P_t} \ln \frac{P_i}{1-P_t}$ <br>Consequentl

$$
H(S_1) = -\sum_{i=0}^{t} \frac{P_i}{P_t} \ln \frac{P_i}{P_t}
$$
 (3)

$$
H(S_2) = -\sum_{i=t+1}^{L-1} \frac{P_i}{1 - P_t} \ln \frac{P_i}{1 - P_t}
$$
 (4)

$$
H(S) = H(S1) + H(S2)
$$
\n
$$
(5)
$$

*H(S<sub>1</sub>)* =  $-\sum_{i=0}^{L} \frac{P_i}{P_i} \ln \frac{P_i}{P_i}$  (3)<br>  $H(S_1) = -\sum_{i=0}^{L} \frac{P_i}{P_i} \ln \frac{P_i}{P_i}$  (3)<br>  $\frac{P_i}{P_i} = -\sum_{i=l+1}^{L-1} \frac{P_i}{1-P_i} \ln \frac{P_i}{1-P_i}$  (4)<br>  $H(S) = H(S_i) + H(S_2)$  (5)<br>  $H(S) = H(S_i) + H(S_2)$  (5)<br>  $H(S) = H(S_i) + H(S_2)$  (5)<br>  $H(S) = H$ currence. Then, let  $P_t = P_0 + P_1 + \dots + P_t$ , the entropy of  $S_t$ <br>
d  $S_2$  can be articulated as:<br>  $H(S_1) = -\sum_{i=0}^t \frac{P_i}{P_i} \ln \frac{P_i}{P_t}$  (3)<br>  $H(S_2) = -\sum_{i=t+1}^{t-1} \frac{P_i}{1-P_t} \ln \frac{P_i}{1-P_t}$  (4)<br>
Consequently, the overall entropy and  $S_2$  can be articulated as:<br>  $H(S_1) = -\sum_{i=0}^{t} \frac{P_i}{P_i} \ln \frac{P_i}{P_t}$  (3)<br>  $H(S_2) = -\sum_{i=t+1}^{t-1} \frac{P_i}{1-P_t} \ln \frac{P_i}{1-P_t}$  (4)<br>
Consequently, the overall entropy of the image can be<br>
expressed as the cumulative of the tw  $H(S_1) = -\sum_{i=0}^{t} \frac{P_i}{P_i} \ln \frac{P_i}{P_t}$  (3)<br>  $H(S_2) = -\sum_{i=t+1}^{t-1} \frac{P_i}{1-P_i} \ln \frac{P_i}{1-P_t}$  (4)<br>
Consequently, the overall entropy of the image can be<br>
expressed as the cumulative of the two types of entropy,<br>
denoted as:<br>  $H(S_1) = -\sum_{i=0}^{n} \frac{P_i}{P_i} \ln \frac{P_i}{P_i}$  (3)<br>  $H(S_2) = -\sum_{i=+1}^{L-1} \frac{P_i}{1-P_i} \ln \frac{P_i}{1-P_i}$  (4)<br>
Consequently, the overall entropy of the image can be<br>
expressed as the cumulative of the two types of entropy,<br>
denoted as:<br>  $\frac{F_i}{f_i} = P_i$   $H_i$  (4)<br>  $H(S_2) = -\sum_{i=1}^{L-1} \frac{P_i}{1-P_i} \ln \frac{P_i}{1-P_i}$  (4)<br>
Consequently, the overall entropy of the image can be<br>
expressed as the cumulative of the two types of entropy,<br>
denoted as:<br>  $H(S) = H(S_i) + H(S_2)$  (5)<br> as: *Theory*  $T_{\text{max}} = \frac{P}{P_{\text{max}}} - \frac{P}{P_{\text{max}}}$  *T*  $\frac{P}{P_{\text{max}}}$  *T*  $\frac{P}{P_{\text{max}}}$  *T*  $\frac{P}{P_{\text{max}}}$  *T Z*  $\frac{P}{P_{\text{max}}}$  *T T Z*  $\frac{P}{P_{\text{max}}}$  *T T Z* (2) that the probability of each grayscale level to *T* pressed as the cumulative of the two types of entropy,<br>noted as:<br> $H(S) = H(S_1) + H(S_2)$  (5)<br>Next, iterate over all grayscale levels t within the range of<br>grayscale histogram, calculating the sum of entropies for<br>expressed is t threshold as:<br>  $H(S) = H(St) + H(Sz)$  (5)<br>
Next, iterate over all grayscale levels t within the range of<br>
the grayscale histogram, calculating the sum of entropies for<br>
the two classes. Then, the grayscale level corresponding to  $H(S) = H(S_1) + H(S_2)$  (5)<br>Next, iterate over all grayscale levels t within the range of<br>the grayscale histogram, calculating the sum of entropies for<br>the two classes. Then, the grayscale level corresponding to<br>the maximum val

$$
T_g = k \cdot |T_{\text{max}} - T_{\text{min}}| \tag{6}
$$

Next, iterate over all grayscale levels t within the range of<br>the grayscale histogram, calculating the sum of entropies for<br>the two classes. Then, the grayscale level corresponding to<br>the maximum value is denoted as  $T_{max}$ the grayscale histogram, calculating the sum of entropies for<br>the grayscale histogram, calculating the sum of entropies for<br>the two classes. Then, the grayscale level corresponding to<br>the maximum value is denoted as  $T_{max}$ the two classes. Then, the grayscale level corresponding to<br>the maximum value is denoted as  $T_{max}$  and the grayscale level<br>corresponding to the maximum value is denoted as  $T_{min}$ .<br>Therefore, the global optimal threshold the maximum value is denoted as  $T_{max}$  and the grayscale level<br>corresponding to the minimum value is denoted as  $T_{min}$ .<br>Therefore, the global optimal threshold  $T_g$  can be represented<br>as:<br> $T_g = k \cdot |T_{max} - T_{min}|$  (6)<br>Where k re For maintain calculate and the corresponding to the minimum value is de<br>
Therefore, the global optimal threshold  $T_g$  can<br>
as:<br>  $T_g = k \cdot |T_{\text{max}} - T_{\text{min}}|$ <br>
Where k represents the scaling factor. Althor<br>
threshold  $T_g$  can Exercise and optimal threshold  $T_g$  can be represented<br>  $T_g = k \cdot |T_{\text{max}} - T_{\text{min}}|$  (6)<br>
Where k represents the scaling factor. Although the global<br>
reshold  $T_g$  can be adaptively selected based on the<br>
stribution of graysc as:<br>  $T_g = k \cdot |T_{\text{max}} - T_{\text{min}}|$  (6)<br>
Where *k* represents the scaling factor. Although the global<br>
threshold *Tg* can be adaptively selected based on the<br>
distribution of grayscale values in the image, it assumes that<br>
all  $T_g = k \cdot |T_{\text{max}} - T_{\text{min}}|$  (6)<br>Where *k* represents the scaling factor. Although the global<br>threshold *Tg* can be adaptively selected based on the<br>distribution of grayscale values in the image, it assumes that<br>all regions Feature *x* Ferricular 1  $I_g = K \cdot |I_{\text{max}} - I_{\text{min}}|$  (6)<br>
Where *k* represents the scaling factor. Although the global<br>
threshold *Tg* can be adaptively selected based on the<br>
distribution of grayscale values in the image, Where *k* represents the scaling factor. Although the global<br>threshold *Tg* can be adaptively selected based on the<br>distribution of grayscale values in the image, it assumes that<br>all regions of the image are under the sam threshold *Tg* can be adaptively selected based on the distribution of grayscale values in the image, it assumes that all regions of the image are under the same lighting conditions. When local regions of the image have v I grayscale levels t within the range of<br>
n, calculating the sum of entropies for<br>
the grayscale level corresponding to<br>
denoted as  $T_{max}$  and the grayscale level<br>
minimum value is denoted as  $T_{min}$ <br>
pptimal threshold  $T_g$ culating the sum of entropies for<br>grayscale level corresponding to<br>ed as  $T_{max}$  and the grayscale level<br>mum value is denoted as  $T_{min}$ .<br>I threshold  $T_g$  can be represented<br>aling factor. Although the global<br>aling factor. A ses. Then, the grayscale level corresponding to<br>ses. Then, the grayscale level corresponding to<br>m value is denoted as  $T_{max}$  and the grayscale level<br>g to the minimum value is denoted as  $T_{min}$ <br>eglobal optimal threshold  $T$ esents the probability of each grayscale level<br>
en, let  $P_t = P_0 + P_1 + \cdots + P_t$ , the entropy of S,<br>
ritculated as:<br>  $H(S_1) = -\sum_{i=0}^{t} \frac{P_i}{P_i} \ln \frac{P_i}{P_i}$  (3)<br>  $f(S_2) = -\sum_{i=0}^{t} \frac{P_i}{P_i} \ln \frac{P_i}{P_i}$  (4)<br>  $f(S_2) = -\sum_{i=t+1}^{t}$ 

Therefore, in situations where lighting conditions vary<br>significantly, our study adopts a local adaptive threshold  $T_i$  to<br>address this issue. Assuming point  $A$  ( $x_0$ ,  $y_0$ ) is a potential<br>feature point in the image. A *I***i** and the grays are average value of region *N*, centered *A*, is selected with a side length of L. Then, the local adaptive threshold  $T_l$  can be represented as:<br>  $T_l = k \cdot \frac{\left[\frac{1}{n} \sum_{i=1}^{n} I_{i_{\text{max}}} - \frac{1}{m} \sum_{i=1}$ 

$$
T_{l} = k \cdot \frac{\left[ \frac{1}{n} \sum_{i=1}^{n} I_{i_{\max}} - \frac{1}{m} \sum_{i=1}^{m} I_{i_{\min}} \right]}{I_{i_{\max}}} \tag{7}
$$

scaling from Sole. It solely<br>feature point in the image. A square region N, centered at A,<br>is selected with a side length of L. Then, the local adaptive<br>threshold  $T_l$  can be represented as:<br> $T_l = k \cdot \frac{\left[\frac{1}{n} \sum_{i=1}^{n} I_{$ selected with a side length of L. Then, the local adaptive<br>reshold  $T_l$  can be represented as:<br> $T_l = k \cdot \frac{\left[ \frac{1}{n} \sum_{i=1}^{n} I_{i_{\text{max}}} - \frac{1}{m} \sum_{i=1}^{m} I_{i_{\text{min}}} \right]}{I_{i_{\text{ave}}}$  (7)<br>In equation (7), we define the maximum gra

reshold  $T_l$  can be represented as:<br>  $T_l = k \cdot \frac{\left[\frac{1}{n} \sum_{i=1}^{n} I_{i_{\text{max}}} - \frac{1}{m} \sum_{i=1}^{m} I_{i_{\text{min}}} \right]}{I_{i_{\text{ave}}}}$  (7)<br>
In equation (7), we define the maximum gray value as  $I_{i_{\text{max}}}$ <br>
d the minimum gray value as  $I_{i_{$  $T_i = k \cdot \frac{\left[\frac{1}{n} \sum_{i=1}^{n} I_{i_{\text{max}}} - \frac{1}{m} \sum_{i=1}^{m} I_{i_{\text{min}}} \right]}{I_{i_{\text{ave}}}$  (7)<br>
In equation (7), we define the maximum gray value as  $I_{i_{\text{max}}}$ <br>
and the minimum gray value as  $I_{min}$  in region N. Moreover,<br>  $I_{i_{\text{ave}}}$  $T_i = k \cdot \frac{\left[ \frac{1}{n} \sum_{i=1}^{n} I_{i_{\text{max}}} - \frac{1}{m} \sum_{i=1}^{m} I_{i_{\text{min}}} \right]}{I_{i_{\text{max}}}$  (7)<br>In equation (7), we define the maximum gray value as  $I_{\text{max}}$ <br>and the minimum gray value as  $I_{\text{min}}$  in region N. Moreover,<br> $I_{\text{layer}}$  d  $T_i = k \cdot \frac{\boxed{n}{2} \cdot \boxed{n_{i_{\text{max}}}} - \boxed{n}{2} \cdot \boxed{l_{i_{\text{min}}}}$  (7)<br>
In equation (7), we define the maximum gray value as  $I_{\text{max}}$ <br>
and the minimum gray value as  $I_{\text{min}}$  in region N. Moreover,<br>  $I_{\text{layer}}$  denotes the grayscale aver  $T_i = k \cdot \frac{N_i - N_i}{N_{i_{\text{user}}}}$  (1)<br>
In equation (7), we define the maximum gray value as  $I_{\text{linear}}$ <br>
and the minimum gray value as  $I_{\text{min}}$  in region N. Moreover,<br>  $I_{\text{layer}}$  denotes the grayscale average value of region N. T

**Volume 32, Issue 10, October 2024, Pages 1909-1920**

**Enginee**<br>and extract semantic information.<br>YOLOv7 algorithm, as a typical representative cone-Stage target detection algorithm [26], has good real-tim<br>performance, and its framework architecture is illustrated in **Engineering Letters**<br>
dextract semantic information.<br>
YOLOv7 algorithm, as a typical representative of introduce supplementary parameter-<br>
Stage target detection algorithm [26], has good real-time the module is that major **Engineering Letters**<br>and extract semantic information.<br>YOLOv7 algorithm, as a typical representative of introduce supplementary pone-Stage target detection algorithm [26], has good real-time the module is that majori<br>perf **Engineering Letters**<br>
and extract semantic information.<br>
YOLOv7 algorithm, as a typical representative of introduce supplementary paramet<br>
One-Stage target detection algorithm [26], has good real-time the module is that m **Engineering Letters**<br>
and extract semantic information. attention weights 1<br>
YOLOv7 algorithm, as a typical representative of introduce supplemer<br>
One-Stage target detection algorithm [26], has good real-time the module i **Engineering Letters**<br>
Manustract semantic information.<br>
TOLOV7 algorithm, as a typical representative of introduce supelementary parameter<br>
One-Stage target detection algorithm [26], has good real-time the module is that **Engineering Letters**<br>
and extract semantic information.<br>
YOLOv7 algorithm, as a typical representative of introduce supplementary pa<br>
One-Stage target detection algorithm [26], has good real-time the module is that majori **Engineering Letters**<br>
and extract semantic information.<br>
YOLOv7 algorithm, as a typical representative of introduce supplementary para<br>
One-Stage target detection algorithm [26], has good real-time the module is that majo **Engineering Letters**<br>
and extract semantic information. attention weights for feature<br>
YOLOv7 algorithm, as a typical representative of introduce supplementary parar<br>
One-Stage target detection algorithm [26], has good re Extended-ELAN), which utilizes extension, stochastic and extract semantic information.<br>
Mone-Stage target detection algorithm [26], has good real-time the module is that major<br>
performance, and its framework architecture i and extract semantic information. The action and extract semantic information.<br>
MCLOv7 algorithm, as a typical representative of introduce supplementary<br>
One-Stage target detection algorithm [26], has good real-time the mo and extract semantic information.<br>
YOLOv7 algorithm, as a typical representative of introduce supplementary para-<br>
One-Stage target detection algorithm [26], has good real-time<br>
the module is that majority of<br>
performance, and extract semantic information. The act of the module strate integrity one-Stage target detection algorithm [26], has good real-time the module is that major performance, and its framework architecture is illustrated in YOLOv7 algorithm, as a typical representative of introduce supplementary p<br>
One-Stage target detection algorithm [26], has good real-time the module is that majority<br>
performance, and its framework architecture is illustra One-Stage target detection algorithm [26], has good real-time the module is that major performance, and its framework structure is illustrated in on the selection of special (RepConv) [27] in the network structure, which d performance, and its framework architecture is illustrated in on the selection of specify (Fig. 6. First, YOLOv7 introduces reparameterization necessity for structural all (RepConv) [27] in the network structure, which dec Fig. 6. First, YOLOv7 introduces reparameterization necessity for structural alteration (RepConv) [27] in the network structure, which decreases the features:<br>
parameter count and computational demands, thereby 1) SimAM is (RepConv) [27] in the network structure, which decreases the features:<br>
parameter count and computational demands, thereby 1) SimAM is able to model r<br>
improving the network's operational efficiency. Then, at multiple lev parameter count and computational demands, thereby 1) SimAM is able to model rel improving the network's operational efficiency. Then, at multiple levels, and it correct CIAN), which utilizes extension, stochastic model to improving the network's operational efficiency. Then, at multiple levels, and it (YOLOv7) improved ELAN by proposing E-ELAN features and high-level server (Extended-ELAN), which utilizes extension, stochastic model to unde YOLOv7 improved ELAN by proposing E-ELAN features and high-le<br>
(Extended-ELAN), which utilizes extension, stochastic model to under<br>
disruption, and merging bases to achieve continuous comprehensively.<br>
enhancement of the xtended-ELAN), which utilizes extension, stochastic model to understand<br>
struption, and merging bases to achieve continuous comprehensively.<br>
SimAM not only focuses on the seserving the integrity of the original gradient p disruption, and merging bases to achieve continuous<br>
emhancement of the network's learning capabilities while<br>
2) SimAM not only for<br>
emerging the integrity of the original gradient pathways.<br>
Musconsity is also able<br>
also enhancement of the network's learning capabilities while 2) SimAM not only focuses on<br>preserving the integrity of the original gradient pathways. but is also able to mine<br>Also, E-ELAN can instruct different computational So, E-ELAN can instruct different computational modules<br>
learn more different features. The label assignment<br>
thodology employed by YOLOv7 integrates the cross-grid<br>
the representational capital<br>
anche technique utilized i to learn more different features. The label assignment<br>
methodology employed by YOLOv7 integrates the cross-grid<br>
search technique utilized in YOLOv5, along with the<br>
matching strategy adopted in YOLOv5. In addition to thi

precision.

methodology employed by YOLOv7 integrates the cross-grid<br>
search technique utilized in YOLOv5, along with the<br>
training strategy adopted in YOLOx. In addition to this, the<br>
training approach incorporating an auxiliary head search technique utilized in YOLOv5, along with the representational capabilities<br>
matching strategy adopted in YOLOx. In addition to this, the 3) SimAM introduces spatial<br>
training approach incorporating an auxiliary head matching strategy adopted in YOLOx. In addition to this, the 3) SimAM introduces spatial is training approach incorporating an auxiliary head is used in the spatial distribution of YOLOv7, which enhances the accuracy by el training approach incorporating an auxiliary head is used in<br>
YOLOv7, which enhances the accuracy by elevating the<br>
training expenses and does not affect the inference time.<br>
Although YOLOv7 runs faster, its detection accu YOLOv7, which enhances the accuracy by elevating the<br>training expenses and does not affect the inference time.<br>
localize the target region,<br>
Although YOLOv7 runs faster, its detection accuracy is<br>
lottention Mechanism) [2 training expenses and does not affect the inference time. localize the target<br>
Although YOLOv7 runs faster, its detection accuracy is<br>
traget detection and<br>
lower, so this paper introduces SimAM (Spatial information 4) Sim Although YOLOv7 runs faster, its detection accuracy is<br>
lower, so this paper introduces SimAM (Spatial information<br>
4) SimAM is adaptive in that<br>
Attention Mechanism) [28] in YOLOv7 to enhance its<br>
orderateristics of atten lower, so this paper introduces SimAM (Spatial information 4) SimAM is adaptive in Attention Mechanism) [28] in YOLOv7 to enhance its digits its focus of atter<br>precision. <br>
The Attention Mechanism is a technique widely emp Attention Mechanism) [28] in YOLOv7 to enhance its<br>
generation.<br>
The Attention Mechanism is a technique widely employed<br>
in computer science and machine learning, where the output<br>
of each neuron is influenced not only by precision.<br>
The Attention Mechanism is a technique widely employed<br>
in computer science and machine learning, where the output<br>
of each neuron is influenced not only by the outputs of all<br>
neurons in the preceding layer b

g Letters<br>attention weights for feature maps without the need to<br>introduce supplementary parameters. An additional benefit of<br>the module is that majority of the operations are predicated<br>on the selection of specified energ g Letters<br>attention weights for feature maps without the need to<br>introduce supplementary parameters. An additional benefit of<br>the module is that majority of the operations are predicated<br>on the selection of specified energ **Example 18 Example 10**<br>attention weights for feature maps without the need to<br>introduce supplementary parameters. An additional benefit of<br>the module is that majority of the operations are predicated<br>on the selection of s **Exercise 15 Exercise 15 Exercise 2016**<br> **Exercise 2 g Letters**<br>attention weights for feature maps without the need to<br>introduce supplementary parameters. An additional benefit of<br>the module is that majority of the operations are predicated<br>on the selection of specified ene features: **Exercise 1)**<br> **Exercise 1)**<br> **Exercise 1)**<br> **Exercise 1)**<br> **EXERCIST AM is that majority of the operations are predicated<br>
on the selection of specified energy functions, reducing the<br>
necessity for structural alterations** etters<br>
attion weights for feature maps without the need to<br>
dduce supplementary parameters. An additional benefit of<br>
module is that majority of the operations are predicated<br>
he selection of specified energy functions, r **Exercise and High-Level Servel Servel Server Server Server Server Server Server School Server School Server School Server School Server School Server School Server Server Server Server Server Server Server Server Server S Exercise 19.1**<br> **Exercise 10.1**<br> attention weights for feature maps without the need to<br>introduce supplementary parameters. An additional benefit of<br>the module is that majority of the operations are predicated<br>on the selection of specified energy function ition weights for feature maps without the need to<br>duce supplementary parameters. An additional benefit of<br>module is that majority of the operations are predicated<br>he selection of specified energy functions, reducing the<br>s

- comprehensively.
- duce supplementary parameters. An additional benefit of<br>module is that majority of the operations are predicated<br>he selection of specified energy functions, reducing the<br>ssity for structural alterations. SimAM has the foll module is that majority of the operations are predicated<br>he selection of specified energy functions, reducing the<br>ssity for structural alterations. SimAM has the following<br>tres:<br>SimAM is able to model relationships between he selection of specified energy functions, reducing the<br>ssity for structural alterations. SimAM has the following<br>tres:<br>SimAM is able to model relationships between features<br>at multiple levels, and it can focus on both lo ssity for structural alterations. SimAM has the followin<br>res:<br>SimAM is able to model relationships between feature<br>at multiple levels, and it can focus on both low-leve<br>features and high-level semantic features, enabling t features:<br>
1) SimAM is able to model relationships between features<br>
at multiple levels, and it can focus on both low-level<br>
features and high-level semantic features, enabling the<br>
model to understand the input data more<br> SimAM is able to model relationships between features<br>at multiple levels, and it can focus on both low-level<br>features and high-level semantic features, enabling the<br>model to understand the input data more<br>comprehensively.<br> at multiple levels, and it can focus on both low-level<br>features and high-level semantic features, enabling the<br>model to understand the input data more<br>comprehensively.<br>SimAM not only focuses on the features of each channel features and high-level semantic features, enabling the<br>model to understand the input data more<br>comprehensively.<br>SimAM not only focuses on the features of each channel,<br>but is also able to mine the relationships between<br>di model to understand the input data more<br>comprehensively.<br>SimAM not only focuses on the features of each channel,<br>but is also able to mine the relationships between<br>different channels. This attention mechanism enables<br>SimAM comprehensively.<br>
2) SimAM not only focuses on the features of each channel,<br>
but is also able to mine the relationships between<br>
different channels. This attention mechanism enables<br>
SimAM to better learn the semantic con SimAM not only focuses on the features of each channel,<br>but is also able to mine the relationships between<br>different channels. This attention mechanism enables<br>SimAM to better learn the semantic connections<br>between feature
- 
- but is also able to mine the relationships between<br>different channels. This attention mechanism enables<br>SimAM to better learn the semantic connections<br>between features and improves the model's<br>representational capabilities different channels. This attention mechanism enables<br>SimAM to better learn the semantic connections<br>between features and improves the model's<br>representational capabilities.<br>SimAM introduces spatial information and focuses SimAM to better learn the semantic connections<br>between features and improves the model's<br>representational capabilities.<br>SimAM introduces spatial information and focuses on<br>the spatial distribution of features. This attenti between features and improves the model's<br>representational capabilities.<br>SimAM introduces spatial information and focuses on<br>the spatial distribution of features. This attention<br>mechanism enhances the model's ability to pr representational capabilities.<br>
SimAM introduces spatial information and focuses on<br>
the spatial distribution of features. This attention<br>
mechanism enhances the model's ability to precisely<br>
localize the target region, in SimAM introduces spatial information and focuses on<br>the spatial distribution of features. This attention<br>mechanism enhances the model's ability to precisely<br>localize the target region, increasing the precision of<br>target de the spatial distribution of features. This attention<br>mechanism enhances the model's ability to precisely<br>localize the target region, increasing the precision of<br>target detection and segmentation.<br>SimAM is adaptive in that mechanism enhances the model's ability to precisely<br>localize the target region, increasing the precision of<br>target detection and segmentation.<br>4) SimAM is adaptive in that it is able to automatically<br>adjust its focus of at localize the target region, increasing the precision of<br>target detection and segmentation.<br>4) SimAM is adaptive in that it is able to automatically<br>adjust its focus of attention according to the different<br>characteristics o target detection and segmentation.<br>
4) SimAM is adaptive in that it is able to automatically<br>
adjust its focus of attention according to the different<br>
characteristics of the input data. This adaptability allows<br>
SimAM to SimAM is adaptive in that it is able to automatically<br>adjust its focus of attention according to the different<br>characteristics of the input data. This adaptability allows<br>SimAM to be applied to a variety of different type adjust its focus of attention according to the different<br>characteristics of the input data. This adaptability allows<br>SimAM to be applied to a variety of different types of<br>data, improving the module's capacity for general characteristics of the input data. This adaptability allows<br>
SimAM to be applied to a variety of different types of<br>
data, improving the module's capacity for generalization.<br>
As illustrated in Fig. 7, SimAM estimates<br>
th



**Volume 32, Issue 10, October 2024, Pages 1909-1920**











**E.** *Target Detection Combined with Eippolar Line*<br> *E. Target Detection Combined with Eippolar Line*<br> *Constraint to Reject Dynamic Feature Points*<br>
Previous SLAM methods based on target detection relied<br>
celusively on t **E.** Target Detection Combined with Eippolar Line<br> *Constraint to Reject Dynamic Feature Points*<br>
Previous SLAM methods based on target detection relied<br>
celusively on these results to eliminate dynamic feature<br>
pints. How **Example 1.1** Engineering Letters<br>
E. Target Detection Combined with Eippolar Line<br>
Constraint to Reject Dynamic Feature Points<br>
Previous SLAM methods based on target detection relied<br>
clusively on these results to elimin **Example 1.1** Entertain whether inherently mobile objects, such as carrenains<br>
External to Reject Dynamic Feature Points<br>
Previous SLAM methods based on target detection relied<br>
exclusively on these results to eliminate d **Example 1.1** Englineering Letters<br>
E. Target Detection Combined with Eippolar Line<br>
Constraint to Reject Dynamic Feature Points<br>
Previous SLAM methods based on target detection relied<br>
exclusively on these results to eli **Example Example Example 2**<br> **E.** Target Detection Combined with Eippolar Line Then, we express the distance<br>
Constraint to Reject Dynamic Feature Points<br>
Previous SLAM methods based on target detection relied<br>
exclusivel **E.** Target Detection Combined with Eippolar Line<br>
Then, we express the dis<br>
Constraint to Reject Dynamic Feature Points<br>
Then, we express the dis<br>
current of the exclusively on these results to eliminate dynamic feature<br> **E.** Target Detection Combined with Eippolar Line<br>
Constraint to Reject Dynamic Feature Points<br>
ine as d, which is a<br>
calculated as follows:<br>
SEAM methods based on target detection relied<br>
exclusively on these results to **E.** Target Detection Combined with Eippolar Line<br>
Constraint to Reject Dynamic Feature Points<br>
Then, we express the distant<br>
Constraint to Reject Dynamic Feature Points<br>
Previous SLAM methods based on target detection re E. Target Detection Combined with Eippolar Line<br>
Constraint to Reject Dynamic Feature Points<br>
Previous SLAM methods based on target detection relied<br>
exclusively on these results to eliminate dynamic feature<br>
points. Howe *E. Target Detection Combined with Eippolar Line* Then, we express the distant Constraint to Reject Dynamic Feature Points in the and  $d$  exclusively on these results to eliminate dynamic feature points. However, target d *E. Target Detection Combined with Eippolar Line*<br>
Constraint to Reject Dynamic Feature Points<br>
Previous SLAM methods based on target detection relied<br>
exclusively on these results to eliminate dynamic feature<br>
points. Ho *E. Target Detection Combined with Eippolar Line*<br>
In as d, which is also known<br>
Constraint to Reject Dynamic Feature Points<br>
exclusively on these results to eliminate dynamic feature<br>
points. However, target detection ne Constraint to Reject Dynamic Feature Points<br>
calculated as fo<br>
exclusively on these results to eliminate dynamic feature<br>
points. However, target detection networks often struggle to<br>
ascertain whether inherently mobile o Previous SLAM methods based on target detection relied<br>
clusively on these results to eliminate dynamic feature<br>
ints. However, target detection networks often struggle to<br>
creatin whether inherently mobile objects, such exclusively on these results to eliminate dynamic feature<br>
points. However, target detection networks often struggle to<br>
ascertain whether inherently mobile objects, such as cars, are<br>
extremely in motion. Consequently, e points. However, target detection networks often struggle to a<br>secretain whether inherently mobile objects, such as cars, are<br>currently in motion. Consequently, even when a car remains the prom Eq. 11, in an ideal sect<br>at ascertain whether inherently mobile objects, such as cars, are<br>currently in motion. Consequently, even when a car remains<br>discard associated feature points, markedly reducing the point *p*<sub>3</sub>, representing the current<br>dis currently in motion. Consequently, even when a car remains<br>stationary in the environment, the system may erroneously<br>onit  $p_2$ , representing the<br>discard associated feature points, markedly reducing the pool<br>freepore, it

stationary in the environment, the system may erroneously<br>
discard associated feature points, markedly reducing the pool<br>
Therefore, it may be considered<br>
available for pose estimation. Moreover, these approaches<br>
frequen discard associated feature points, markedly reducing the pool<br>available for pose estimation. Moreover, these approaches<br>frequently fail to filter out feature points attributed to static moise typically results in d<br>points available for pose estimation. Moreover, these approaches<br>frequently fail to filter out feature points attributed to static<br>objects like books or chairs that are being displaced by<br>individuals, resulting in inaccurate data Expectively fail to filter out feature points attributed to static<br>
income typically results in the<br>
discussion charged and in this paper, the threshold is<br>
finding in inaccurate data association in It his paper, the thre objects like books or chairs that are being displaced by<br>
individuals, resulting in inaccurate data associations and a<br>
in this paper, the threshold is<br>
significant degradation in SLAM system precision.<br>
In order to tackl individuals, resulting in inaccurate data associations and a<br>
significant degradation in SLAM system precision.<br>
Ins order to tackle these challenges, we integrate target<br>
To summarize, the step detection with an epipolar significant degradation in SLAM system precision. <br>
less than e, we regard then order to tackle these challenges, we integrate target To summarize, the detection with an epipolar constraint approach to eliminate points in

In order to tackle these challenges, we integrate target To summarize, the steps for<br>detection with an epipolar constraint approach to eliminate points in our study are as follow<br>dfp. First, we need to align the feature p detection with an epipolar constraint approach to eliminate points in our study are as follows<br>dfp. First, we need to align the feature points from two Step 1: We employ the implomental matrix. Next, we can assess the dis consecutive frames and use them to calculate the various targets in debli<br>fundamental matrix. Next, we can assess the distance positions within these<br>between the feature points in the current frame and their detection box fundamental matrix. Next, we can assess the distance positions within<br>between the feature points in the current frame and their detection boxes. In<br>related epipolar lines. Greeater distances indicate a higher about the ta



as:

$$
P_1 = [x_1, y_1, 1], P_2 = [x_2, y_2, 1]
$$
 (8)

The chi-square coordinates of 
$$
P_1
$$
 and  $P_2$  can be expressed  
\nas:  
\n $P_1 = [x_1, y_1, 1], P_2 = [x_2, y_2, 1]$   
\nIn which x and y represent the pixel coordinate values of  
\nthe feature points, the polar line  $l_2$  can be expressed as:  
\n
$$
l_2 = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = F \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}
$$
  
\n
$$
l_3 = \begin{bmatrix} 80 & 1000 \\ 1 & 1 \end{bmatrix}
$$
  
\n
$$
l_4 = \begin{bmatrix} 1 & 000 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_5 = \begin{bmatrix} 1 & 000 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_6 = \begin{bmatrix} 1 & 000 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_7 = \begin{bmatrix} 1 & 000 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_8 = \begin{bmatrix} 1 & 000 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_9 = \begin{bmatrix} 1 & 00 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{bmatrix} 1 & 00 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{bmatrix} 1 & 00 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}
$$
  
\n
$$
l_1 = \begin{
$$

$$
p_2^T F p_1 = p_2^T l_2 = 0 \tag{10}
$$

**Then, we express the distance between** *p***<sub>2</sub> and the epipolar e as** *d***, which is also known as the offset dist. It is levelated as follows:<br>** $\begin{bmatrix} p_2^T F p_1 \end{bmatrix}$  **(11) EXECUTE:**<br>
Then, we express the distance between  $p_2$  and the epipolar<br>
line as *d*, which is also known as the offset dist. It is<br>
calculated as follows:<br>  $d = \frac{|p_2^T F p_1|}{\sqrt{x^2 + x^2}}$  (11) **g Letters**<br>Then, we express the distance between  $p_2$  and the<br>line as d, which is also known as the offset d<br>calculated as follows:<br> $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$ ce between  $p_2$  and the epipolar<br>
own as the offset dist. It is<br>  $\frac{p_2^T F p_1}{r^2 + Y^2}$  (11)<br>
enario, when  $d = 0$ , the feature<br>
rrent frame, is situated on  $l_2$ .

$$
d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}
$$
 (11)

2 2 Figure 1.1 and the epipolar<br> **Theorem** as the offset dist. It is<br>  $\frac{p_2^T F p_1}{X^2 + Y^2}$ (11)<br>  $\frac{X^2 + Y^2}{X^2 + Y^2}$ (11)<br>
cenario, when  $d = 0$ , the feature<br>
urrent frame, is situated on  $l_2$ <br>
red static.<br>
presence of v **Example 11, 11, 11, in an ideal scenario,** the feature of the feature of the feature data scenario, when  $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>From Eq. 11, in an ideal scenario, when  $d = 0$ , the feature int  $p_2$ , represent **pointiff**<br> **point 1.1.** Then, we express the distance between  $p_2$  and the epipolar<br>
line as *d*, which is also known as the offset dist. It is<br>
calculated as follows:<br>  $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>
From Eq. 11, i

ance between  $p_2$  and the epipolar<br>
mown as the offset dist. It is<br>  $\frac{p_2^T F p_1}{X^2 + Y^2}$ <br>
cenario, when  $d = 0$ , the feature<br>
current frame, is situated on  $l_2$ <br>
red static.<br>
presence of various forms of<br>
ne offset di **Example 12.11**<br>
Then, we express the distance between  $p_2$  and the epipolar<br>
line as d, which is also known as the offset dist. It is<br>
calculated as follows:<br>  $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>
From Eq. 11, in an ideal Then, we express the distance between p<sub>2</sub> and the epipolar<br>
e as d, which is also known as the offset dist. It is<br>
cludated as follows:<br>  $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>
From Eq. 11, in an ideal scenario, when  $d = 0$ , Then, we express the distance between  $p_2$  and the epipolar<br>line as d, which is also known as the offset dist. It is<br>calculated as follows:<br> $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>From Eq. 11, in an ideal scenario, when  $d = 0$ Then, we express the distance between  $p_2$  and the epipolar<br>line as *d*, which is also known as the offset dist. It is<br>calculated as follows:<br> $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>From Eq. 11, in an ideal scenario, when *d* and an external in a static distribution of the inproved YOLOv7 to identify<br>
the point point as static. It is calculated as follows:<br>  $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>
From Eq. 11, in an ideal scenario, when  $d = 0$ , the The steaded as follows.<br>  $d = \frac{|p_2^T F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>
From Eq. 11, in an ideal scenario, when  $d = 0$ , the feature<br>
int  $p_2$ , representing the current frame, is situated on  $l_2$ <br>
lerefore, it may be considered sta  $d = \frac{|p'_2 F p_1|}{\sqrt{X^2 + Y^2}}$  (11)<br>
From Eq. 11, in an ideal scenario, when  $d = 0$ , the feature<br>
point p<sub>2</sub>, representing the current frame, is situated on l<sub>2</sub>.<br>
Therefore, it may be considered static.<br>
However, in practic  $d = \frac{1}{\sqrt{X^2 + Y^2}}$  (11)<br>
From Eq. 11, in an ideal scenario, when  $d = 0$ , the feature<br>
int  $p_2$ , representing the current frame, is situated on  $l_2$ <br>
lerefore, it may be considered static.<br>
However, in practice, the pr

 $\sqrt{X^2 + Y^2}$ <br>
From Eq. 11, in an ideal scenario, when  $d = 0$ , the feature<br>
point  $p_2$ , representing the current frame, is situated on  $l_2$ .<br>
Therefore, it may be considered static.<br>
However, in practice, the presence o From Eq. 11, in an ideal scenario, when  $d = 0$ , the feature<br>point  $p_2$ , representing the current frame, is situated on  $l_2$ .<br>Therefore, it may be considered static.<br>However, in practice, the presence of various forms of point  $p_2$ , representing the current frame, is situated on  $l_2$ .<br>Therefore, it may be considered static.<br>However, in practice, the presence of various forms of<br>noise typically results in the offset dist exceeding 0, yet Therefore, it may be considered static.<br>
However, in practice, the presence of various forms of<br>
noise typically results in the offset dist exceeding 0, yet<br>
remaining below a defined empirical threshold, denoted as  $\varepsilon$ tasks. ise typically results in the offset dist exceeding 0, yet<br>maining below a defined empirical threshold, denoted as  $\varepsilon$ .<br>this paper, the threshold is selected as  $0.6$ , and when  $d$  is<br>s than  $\varepsilon$ , we regard this feature remaining below a defined empirical threshold, denoted as  $\varepsilon$ .<br>In this paper, the threshold is selected as 0.6, and when *d* is<br>less than  $\varepsilon$ , we regard this feature point as static.<br>To summarize, the steps for removin

In this paper, the threshold is selected as 0.6, and when  $d$  is less than  $\varepsilon$ , we regard this feature point as static.<br>To summarize, the steps for removing dynamic feature points in our study are as follows:<br>Step 1: We less than  $\varepsilon$ , we regard this feature point as static.<br>To summarize, the steps for removing dynamic feature<br>points in our study are as follows:<br>Step 1: We employ the improved YOLOv7 to identify<br>various targets in deblurr To summarize, the steps for removing dynamic feature<br>points in our study are as follows:<br>Step 1: We employ the improved YOLOv7 to identify<br>various targets in deblurred images and ascertain their<br>positions within these imag points in our study are as follows:<br>
Step 1: We employ the improved YOLOv<br>
various targets in deblurred images and as<br>
positions within these images, while also<br>
detection boxes. In addition, we extract semanti<br>
about the deblurred images and ascertain their<br>nese images, while also delineating<br>ddition, we extract semantic information<br>provide necessary data for subsequent<br>b determine whether those feature points<br>ject detection frame satisfy In this section, we extract semantic information<br>
we asses. In addition, we extract semantic information<br>
we the targets to provide necessary data for subsequent<br>
ks.<br>
Step 2: We need to determine whether those feature poi detection boxes. In addition, we extract semantic information<br>about the targets to provide necessary data for subsequent<br>tasks.<br>Step 2: We need to determine whether those feature points<br>in the dynamic object detection fram

about the targets to provide necessary data for subsequent<br>tasks.<br>Step 2: We need to determine whether those feature points<br>in the dynamic object detection frame satisfy epipolar<br>constraint. If the feature points in the de tasks.<br>
Step 2: We need to determine whether those feature points<br>
in the dynamic object detection frame satisfy epipolar<br>
constraint. If the feature points in the detection frame don't<br>
satisfy epipolar constraint, they a Step 2: We need to determine whether those teature points<br>in the dynamic object detection frame satisfy epipolar<br>constraint. If the feature points in the detection frame don't<br>satisfy epipolar constraint, they are consider in the dynamic object detection frame satisty epipolar<br>constraint. If the feature points in the detection frame don't<br>satisfy epipolar constraint, they are considered as dynamic<br>feature points and are no longer used in the constrant. If the teature points in the detection frame don't<br>satisfy epipolar constraint, they are considered as dynamic<br>feature points and are no longer used in the subsequent<br>tracking threads.<br>IV. EXPERIMENTS<br>In this se satisty epipolar constraint, they are considered as dynamic<br>feature points and are no longer used in the subsequent<br>tracking threads.<br>IV. EXPERIMENTS<br>In this section, we assess the efficacy of our system by<br>conducting eval teature points and are no longer used in the<br>tracking threads.<br>IV. EXPERIMENTS<br>In this section, we assess the efficacy of or<br>conducting evaluations employing the TU<br>Subsequently, we compare the outcomes with the<br>from ORB-S IV. EXPERIMENTS<br>In this section, we assess the efficacy of our system by<br>anducting evaluations employing the TUM dataset.<br>ubsequently, we compare the outcomes with those obtained<br>om ORB-SLAM3. In addition, this paper also IV. EXPERIMENTS<br>In this section, we assess the efficacy of our system by<br>nducting evaluations employing the TUM dataset.<br>bsequently, we compare the outcomes with those obtained<br>om ORB-SLAM3. In addition, this paper also co In this section, we assess the efficacy of our system by<br>conducting evaluations employing the TUM dataset.<br>Subsequently, we compare the outcomes with those obtained<br>from ORB-SLAM3. In addition, this paper also compares<br>wit conducting evaluations employing the TUM dataset.<br>Subsequently, we compare the outcomes with those obtained<br>from ORB-SLAM3. In addition, this paper also compares<br>with advanced VS algorithms that utilize object detection i

16 B in the moment of material complete the moment of the state in the state of the state in Fig. 10,  $\theta$  and  $\theta$  is the state in Fig. 10,  $\theta$  and  $\theta$  is the state in Fig. 10,  $\theta$  and  $\theta$  is the state in Fig. 10,  $\$ IN BOGB of memory.<br>
In the TUM RGB-D dataset and evaluation<br>
In TUM RGB-D dataset 130<br>
Vision Group at the Technic<br>
represents a large-scale resource<br>
Fig. 10. Epipolar constraint<br>
The chi-square coordinates of P<sub>1</sub> and P *A. TUM dataset and ev*<br>  $\begin{array}{ccc}\nP_1 \\
\hline\n\end{array}$   $\begin{array}{ccc}\nP_2 \\
\hline\n\end{array}$  The TUM RGB-D dataset and ev<br>
The TUM RGB-D dataset and ev<br>
Fig. 10. Epipolar constraint<br>
Fig. 10. Epipolar constraint<br>
Fig. 10. Epipolar constraint Fig. 10. Euphar point in space. *P* and *B* section, we assess the enemic and *M*<sub>2</sub>, *l* and *I* conducting evaluations employing the T<br>
In this section, We compute the outcomes with the subsequently, we compute the outc as inustrated in Fig. 10, *U<sub>N</sub>* and *O<sub>2</sub>* uncal *O<sub>2</sub>* uncal *C* and *C* is the point *P* on *M*<sub>*i*</sub> and *D<sub>2</sub>* is described in the point *P* on *M*<sub>*i*</sub> and *H*<sub>2</sub>, *I* and *H*<sub>2</sub>, *I* and *H*<sub>2</sub>, *I* and *H*<sub>2</sub>, *I* The movement in the movement of a careter out of the point of the point of the point of the point  $M_2$  is described to the point  $P$  on  $M_1$  and  $M_2$ ,  $h$  and  $h$  conducting evaluations emplies<br>  $\kappa$ . The point  $P$  on arison in point in particular to the point P on M<sub>i</sub> and A<sub>2</sub>, 1<sub>i</sub> and 1<sub>2</sub> conducting evaluations emply<br>sequently, we compare the of the point P on M<sub>i</sub> and  $M_2$ , 1<sub>i</sub> and 1<sub>2</sub> conducting evaluations emply<br>sequently, w Fig. 10. Epipolar constraint<br>
The chi-square coordinates of  $P_1$  and  $P_2$  can be expressed fital values of the system's set our system's<br>  $P_1 = [x_1, y_1, 1], P_2 = [x_2, y_2, 1]$  (8) packets include a greater nu<br>
In which x a with advanced VS algorithms that utilize of<br>dynamic surroundings such satisfies and AHV-SLAM<br>all experiments in this paper were conduct<br>system featuring an Intel iS CPU, RTX10<br>is of Ph p and explicit the TUM RGB-D dataset Subsequently, we compare the outcomes with those obtained<br>from ORB-SLAM3. In addition, this paper also compares<br>with advanced VS algorithms that utilize object detection in<br>dynamic surroundings such as AHY-SLAM and RDS-SL From ORB-SLAM3. In addition, this paper also compares<br>with advanced VS algorithms that utilize object detection in<br>dynamic surroundings such as AHY-SLAM and RDS-SLAM.<br>All experiments in this paper were conducted on a compu with advanced VS algorithms that utilize object detection in<br>dynamic surroundings such as AHY-SLAM and RDS-SLAM.<br>All experiments in this paper were conducted on a computer<br>system featuring an Intel i5 CPU, RTX1050Ti GPU, a dynamic surroundings such as AHY-SLAM and RDS-SLAM.<br>All experiments in this paper were conducted on a computer<br>system featuring an Intel i5 CPU, RTX1050Ti GPU, and<br>16GB of memory.<br>A. TUM dataset and evaluation metrics<br>The All experiments in this paper were conducted on a computer<br>system featuring an Intel i5 CPU, RTX1050Ti GPU, and<br>16GB of memory.<br>A. *TUM dataset and evaluation metrics*<br>The TUM RGB-D dataset [30], provided by the Computer<br>V Existen featuring an Intel is CPU, RTX1050Ti GPU, and 16GB of memory.<br>
A. TUM dataset and evaluation metrics<br>
The TUM RGB-D dataset [30], provided by the Computer<br>
Vision Group at the Technical University of Munich,<br>
repr 16GB of memory.<br>
16GB of memory.<br>
16GB of memory.<br>
26 and evaluation metrics<br>
The TUM RGB-D dataset [30], provided by the Computer<br>
Vision Group at the Technical University of Munich,<br>
represents a large-scale resource tha *A. TUM dataset and evaluation metrics*<br>The TUM RGB-D dataset [30], provided by t<br>Vision Group at the Technical University<br>represents a large-scale resource that has set a for<br>evaluating SLAM systems. We utilize<br>packages t 1. TUM dataset and evaluation metrics<br>The TUM RGB-D dataset [30], provided by the Computer<br>sion Group at the Technical University of Munich,<br>presents a large-scale resource that has set a new standard<br>r evaluating SLAM sys The TUM RGB-D dataset [30], provided by the Computer<br>Vision Group at the Technical University of Munich,<br>represents a large-scale resource that has set a new standard<br>for evaluating SLAM systems. We utilized five data<br>pack Vision Group at the Technical University of Munich,<br>represents a large-scale resource that has set a new standard<br>for evaluating SLAM systems. We utilized five data<br>packages that encompass a significant number of dynamic<br>o represents a large-scale resource that has set a new standard<br>for evaluating SLAM systems. We utilized five data<br>packages that encompass a significant number of dynamic<br>objects to test our system's performance, respectivel for evaluating SLAM systems. We utilized five data<br>packages that encompass a significant number of dynamic<br>objects to test our system's performance, respectively<br>fr3/walking/xyz, fr3/walking/half, fr3/walking/static,<br>fr3/w beta the completion of the system's performance of dynamic<br>
bijects to test our system's performance, respectively<br>
3/walking/xyz, fr3/walking/half, fr3/walking/static,<br>
3/walking/rpy, and fr3/sitting/static. The initial f Soly and the mean of the meaning of the meaning of the meaning the meaning the meaning the meaning the meaning the control of the meaning checkets include a greater number of dynamic objects, altered include a greater numb

method we proposed: Absolute Trajectory Error (ATE), which measures the disparity between estimated and ground<br>truth trajectories, and Relative Pose Error (RPE), utilized to fr3/walking/rpy, and fr3/sitting/static. The initial four data<br>packets include a greater number of dynamic objects,<br>whereas the final data packet has a smaller quantity of<br>dynamic objects.<br>We utilize two metrics to assess

**Engineering Letters**<br>
comparative analysis of the experimental data acquired from dynamic environments, the tra<br>
our system against these obtained from ORB-SLAM3, as system closely matches the groun<br>
shown in Table I to I **Engineering Letters**<br>
comparative analysis of the experimental data acquired from dynamic environments, the tra-<br>
our system against these obtained from ORB-SLAM3, as system closely matches the grour<br>
shown in Table I to **Engineering Letters**<br>
comparative analysis of the experimental data acquired from dynamic environments, the<br>
our system against these obtained from ORB-SLAM3, as system closely matches the gr<br>
shown in Table I to III. In **Engineering Letters**<br>
comparative analysis of the experimental data acquired from<br>
dynamic environments, the t<br>
our system against these obtained from ORB-SLAM3, as<br>
system closely matches the group<br>
shown in Table I to I **Engineering Letters**<br>
comparative analysis of the experimental data acquired from<br>
our system against these obtained from ORB-SLAM3, as<br>
system closely match<br>
shown in Table I to III. In these tables, RMSE (Root Mean<br>
the **Engineering Letters**<br>
comparative analysis of the experimental data acquired from dynamic environments, the tive our system closine of the colume of the SCLAM3, as system closely matches the group of the SCLAM3, as syste **Engineering Letters**<br>
comparative analysis of the experimental data acquired from dynamic environments, our system against these obtained from ORB-SLAM3, as system closely matches the shown in Table I to III. In these tab **Engineering Letters**<br>
comparative analysis of the experimental data acquired from dynamic environments, the tra-<br>
our system against these obtained from ORB-SLAM3, as system closely matches the groun<br>
shown in Table I to **Engineering Letters**<br>
comparative analysis of the experimental data acquired from dynamic environments, the trour system against these obtained from ORB-SLAM3, as system closely matches the groum<br>
shown in Table I to III. **Engineering Letters**<br>
comparative analysis of the experimental data acquired from<br>
our system against these obtained from ORB-SLAM3, as<br>
system closely matches the gre<br>
shown in Table I to III. In these tables, RMSE (Root comparative analysis of the experimental data acquired from<br>
our system against these obtained from ORB-SLAM3, as<br>
system closely matches the<br>
shown in Table I to III. In these tables, RMSE (Root Mean<br>
the proposed system comparative analysis of the experimental data acquired from dynamic environment our system against these tobtained from ORB-SLAM3, as system closely match shown in Table I to III. In these tables, RMSE (Root Mean the propo comparative analysis of the experimental data acquired from<br>
dynamic environments, the<br>
our system against these obtained from ORB-SLAM3, as<br>
system closely matches the<br>
shown in Table I to III. In these tables, RMSE (Root our system against these obtained from ORB-SLAM3, as system closely matches the<br>shown in Table I to III. In these tables, RMSE (Root Mean the proposed system der<br>square Error) quantifies the disparity between predicted and shown in Table I to III. In these tables, RMSE (Root Mean<br>
Square Error) quantifies the disparity between predicted and<br>
localization performan<br>
actual values. S.D. (Standard Deviation) gauges the compared it with rece<br>
to Square Error) quantifies the disparity between predicted and<br>actual values. A lower RMSE signifies closer approximation To furth<br>to the true values. S.D. (Standard Deviation) gauges the<br>compared i<br>spread of values within a tual values. A lower RMSE signifies closer approximation To further validate the advation the true values. S.D. (Standard Deviation) gauges the compared it with recently publication inclearing reduced variability. Tables I to the true values. S.D. (Standard Deviation) gauges the compared it with recently publ<br>spread of values within a dataset, with a smaller standard as RDS-SLAM and AHY-SLA<br>deviation indicating reduced variability and enhanc spread of values within a dataset, with a smaller standard<br>deviation indicating reduced variability and enhanced system system exhibits signifi<br>stability. Tables 1 through III demonstrate that our system the five sequences deviation indicating reduced variability and enhanced system system exhibits significantly lot stability. Tables I through III demonstrate that our system the five sequences compained bulgh-dynamic sequences, confirming th stability. Tables I through III demonstrate that our system<br>
outperforms ORB-SLAM3 by over 91% in most demonstrating its superior p<br>
high-dynamic sequences, confirming the superior capabilities. As a crucial com<br>
performan

outperforms ORB-SLAM3 by over 91% in most demonstrating its sup<br>high-dynamic sequences, confirming the superior capabilities. As a crue<br>performance of our system in highly dynamic environments. real-time capability is<br>fle mather of the supervior appearing the superior capabilities. As a crucial compromement of our system in highly dynamic environments.<br>
The performance in low dynamic surroundings is slightly we also measured the runtime of

g Letters<br>
dynamic environments, the trajectory estimated by our<br>
system closely matches the ground truth trajectory. Therefore,<br>
the proposed system demonstrates superior tracking and<br>
localization performance.<br>
To furthe g Letters<br>
dynamic environments, the trajectory estimated by our<br>
system closely matches the ground truth trajectory. Therefore,<br>
the proposed system demonstrates superior tracking and<br>
localization performance.<br>
To furthe **Example 19 Letters**<br>
dynamic environments, the trajectory estimated by our<br>
system closely matches the ground truth trajectory. Therefore,<br>
the proposed system demonstrates superior tracking and<br>
localization performance. **g Letters**<br>dynamic environments, the trajectory estimates<br>system closely matches the ground truth trajectory.<br>the proposed system demonstrates superior trac<br>localization performance.<br>To further validate the advancements o

Letters<br>
namic environments, the trajectory estimated by our<br>
stem closely matches the ground truth trajectory. Therefore,<br>
a proposed system demonstrates superior tracking and<br>
calization performance.<br>
To further validate **g Letters**<br>dynamic environments, the trajectory estimated by our<br>system closely matches the ground truth trajectory. Therefore,<br>the proposed system demonstrates superior tracking and<br>localization performance.<br>To further v **g Letters**<br>dynamic environments, the trajectory estimated by our<br>system closely matches the ground truth trajectory. Therefore,<br>the proposed system demonstrates superior tracking and<br>localization performance.<br>To further v **g Letters**<br>dynamic environments, the trajectory estimated by our<br>system closely matches the ground truth trajectory. Therefore,<br>the proposed system demonstrates superior tracking and<br>localization performance.<br>To further v **g Letters**<br>dynamic environments, the trajectory estimated by our<br>system closely matches the ground truth trajectory. Therefore,<br>the proposed system demonstrates superior tracking and<br>localization performance.<br>To further v dynamic environments, the trajectory estimated by our<br>system closely matches the ground truth trajectory. Therefore,<br>the proposed system demonstrates superior tracking and<br>localization performance.<br>To further validate the dynamic environments, the trajectory estimated by our<br>system closely matches the ground truth trajectory. Therefore,<br>the proposed system demonstrates superior tracking and<br>localization performance.<br>To further validate the dynamic environments, the trajectory estimated by our system closely matches the ground truth trajectory. Therefore, the proposed system demonstrates superior tracking and localization performance. To further validate the dynamic environments, the trajectory estimated by our system closely matches the ground truth trajectory. Therefore, the proposed system demonstrates superior tracking and localization performance.<br>To further validate the system closely matches the ground truth trajectory. Therefore,<br>the proposed system demonstrates superior tracking and<br>localization performance.<br>To further validate the advancements of our system, we<br>compared it with recent e proposed system demonstrates superior tracking and<br>calization performance.<br>To further validate the advancements of our system, we<br>mpared it with recently published SLAM algorithms, such<br>RDS-SLAM and AHY-SLAM. As shown in localization performance.<br>
To further validate the advancements of our system, we<br>
compared it with recently published SLAM algorithms, such<br>
as RDS-SLAM and AHY-SLAM. As shown in Table IV, our<br>
system exhibits significant To further validate the advancements of our system, we compared it with recently published SLAM algorithms, such as RDS-SLAM and AHY-SLAM. As shown in Table IV, our system exhibits significantly lower RMSE and S.D. across compared it with recently published SLAM algorithms, such<br>as RDS-SLAM and AHY-SLAM. As shown in Table IV, our<br>system exhibits significantly lower RMSE and S.D. across<br>the five sequences compared to these algorithms,<br>demons

as RDS-SLAM and AHY-SLAM. As shown in Table IV, our system exhibits significantly lower RMSE and S.D. across the five sequences compared to these algorithms, demonstrating its superior performance and advanced capabilities system exhibits significantly lower RMSE and S.D. across<br>the five sequences compared to these algorithms,<br>demonstrating its superior performance and advanced<br>capabilities. As a crucial component of mobile robotics,<br>real-ti the five sequences compared to these algorithms,<br>demonstrating its superior performance and advanced<br>capabilities. As a crucial component of mobile robotics,<br>real-time capability is essential for SLAM systems. Hence,<br>we al demonstrating its superior performance and advanced capabilities. As a crucial component of mobile robotics, real-time capability is essential for SLAM systems. Hence, we also measured the runtime of the SLAM systems. The capabilities. As a crucial component of mobile<br>real-time capability is essential for SLAM systems<br>we also measured the runtime of the SLAM system<br>runtime of each SLAM system is illustrated in Table<br>As indicated in Table V, moval process inadvertently As indicated in Table V, we can see that our s<br>feature points. exhibits better real-time performance compare<br>te the estimated trajectories and RDS-SLAM and AHY-SLAM, averaging 75ms per t<br>orbserv

Sequence	ORB-SLAM3		Ours		Improvement/%	
	<b>RMSE</b>	S.D.	<b>RMSE</b>	S.D.	<b>RMSE</b>	S.D.
fr3 walking xyz	0.7126	0.3672	0.0163	0.0086	97.71	97.65
fr3 walking half	0.3942	0.2854	0.0207	0.0158	94.47	94.46
fr3 walking static	0.4001	0.0640	0.0072	0.0038	98.2	94.06
fr3 walking rpy	0.4223	0.3321	0.0443	0.0367	89.52	88.96
fr3_sitting_static	0.0075	0.0044	0.0078	0.0045	$-4.3$	$-3.2$
			TABLE II. Results of metric rotational drift (RPE)			
Sequence	ORB-SLAM3		Ours		Improvement/%	
	<b>RMSE</b>	S.D.	<b>RMSE</b>	S.D.	<b>RMSE</b>	S.D.

Sequence	ORB-SLAM3		Ours			Improvement/%	
	<b>RMSE</b>	S.D.	<b>RMSE</b>	S.D.	<b>RMSE</b>	S.D.	
fr3 walking xyz	6.2841	3.4841	0.2928	0.1972	95.34	94.34	
fr3 walking half	6.8735	5.4233	0.4728	0.4598	93.12	91.52	
fr3 walking static	2.7134	2.2098	0.2374	0.2134	91.25	90.34	
fr3_walking_rpy	5.3785	3.4785	0.7142	0.5725	86.72	83.54	
fr3 sitting static	0.1687	0.0087	0.1781	0.0094	$-5.6$	$-8.2$	
TABLE III. Results of metric translational drift (RPE)							
Sequence	ORB-SLAM3		Ours		Improvement/%		
	<b>RMSE</b>	S.D.	<b>RMSE</b>	S.D.	<b>RMSE</b>	S.D.	













TABLE IV. Comparison of absolute trajectory errors between Ours and other similar SLAM methods. (ATE				



Systems<br>
ORB-SALM3<br>
Ours<br>
Average Processi<br>
ORB-SALM3<br>
Ours<br>
AFIY-SLAM<br>
RDS-SLAM<br>
RDS-SLAM<br>
RDS-SLAM<br>
RDS-SLAM<br>
This paper introduces a real-time semantic SLAM system dfp. Experimental<br>
introduce an image enhancement modul network to deblur camera-captured images, thereby<br>
network to define the state of the state enhancing the quality of image frames. Subsequently, we be enhancing the dying of the same the same the same of the same the same of the s Formation of the detection of the detection of the detection<br>
integrate the outputs from<br>
This paper introduces a real-time semantic SLAM system<br>
This paper introduces a real-time semantic SLAM system<br>
integrating an enhan The system of the system. We integrate the outputs from<br>
introduce an enhanced YOLOv7 network. Ini RDS-SLAM<br>
We integrate the outputs<br>
We integrate the outputs<br>
We integrate the outputs<br>
integrating an enhanced YOLOv7 network. Initially, we<br>
introduce an image enhancement module into the some existing SLAM alg<br>
ORB-SLAM V. CONCLUSION<br>
We integrate the outpublic constrain<br>
introduces a real-time semantic SLAM system<br>
dfp. Experimental rest<br>
introduce an image enhancement module into the some existing SLAM introduce<br>
on introduce an image e V. CONCLUSION<br>
We integrate the outputs from with epipolar constraints in ord<br>
integrating an enhanced YOLOv7 network. Initially, we achieves over 90% improvement<br>
introduce an image enhancement module into the some existi V. CONCLUSION<br>
With epipolar constraints<br>
integrating an enhanced YOLOv7 network. Initially, we<br>
introduce an image enhancement module into the some existing SLAM<br>
introduce an image enhancement module into the some existi This paper introduces a real-time semantic SLAM system dfp. Experimental results show integrating an enhanced YOLOv7 network. Initially, we achieves over 90% improvement incroduce an image enhancement module into the some integrating an enhanced YOLOv7 network. Initially, we achieves over 90% improveme<br>introduce an image enhancement module into the some existing SLAM algorithn<br>ORB-SLAM3 framework, leveraging the DeblurGANv2 the enhanced YOL introduce an image enhancement module into the some existing SLAM<br>
ORB-SLAM3 framework, leveraging the DeblurGANv2 the enhanced YOLC<br>
network to deblur camera-captured images, thereby precision while senhancing the quality

actives over the control of the control of the comparison of something and a meaning SLAM algorithms. The processing Time Per Frame(ms)<br>
for the control of the contrast of the difp. Experimental results show Average Processing Time Per Frame(ms)<br>
62<br>
75<br>
75<br>
82<br>
We integrate the outputs from the object detection process<br>
with epipolar constraints in order to eliminate dynamic the<br>
dfp. Experimental results show that our propos **Example 18 and 18**  $\frac{0.2}{75}$ <br>  $\frac{103}{82}$ <br>
We integrate the outputs from the object detection<br>
with epipolar constraints in order to eliminate dy<br>
dfp. Experimental results show that our propos<br>
achieves over 90% improvement in accuracy  $\frac{82}{100}$ <br>  $\frac{82}{100}$ <br>  $\frac{82}{1000}$ <br>  $\frac{82}{1000}$ <br>  $\frac{82}{1000}$ <br>
Experimental results show that our proposed system<br>
hieves over 90% improvement in accuracy compared to<br>
me existing SLAM algorithms. Furthermore, b <sup>103</sup><br>82<br>82<br>82<br>82<br>We integrate the outputs from the object detection process<br>with epipolar constraints in order to eliminate dynamic the<br>dfp. Experimental results show that our proposed system<br>achieves over 90% improvement We integrate the outputs from the object detection process<br>with epipolar constraints in order to eliminate dynamic the<br>dfp. Experimental results show that our proposed system<br>achieves over 90% improvement in accuracy comp We integrate the outputs from the object detection process<br>with epipolar constraints in order to eliminate dynamic the<br>dfp. Experimental results show that our proposed system<br>achieves over 90% improvement in accuracy compa

We integrate the outputs from the object detection process<br>with epipolar constraints in order to eliminate dynamic the<br>dfp. Experimental results show that our proposed system<br>achieves over 90% improvement in accuracy compa with epipolar constraints in order to eliminate dynamic the dfp. Experimental results show that our proposed system achieves over 90% improvement in accuracy compared to some existing SLAM algorithms. Furthermore, by utili dfp. Experimental results show that our proposed system<br>achieves over 90% improvement in accuracy compared to<br>some existing SLAM algorithms. Furthermore, by utilizing<br>the enhanced YOLOv7 algorithm, our system enhances<br>prec achieves over 90% improvement in accuracy compared to<br>some existing SLAM algorithms. Furthermore, by utilizing<br>the enhanced YOLOv7 algorithm, our system enhances<br>precision while simultaneously preserving real-time<br>operatio some existing SLAM algorithms. Furthermore, by utilizing<br>the enhanced YOLOv7 algorithm, our system enhances<br>precision while simultaneously preserving real-time<br>operational efficiency.<br>Although the system proposed in this p

**Engineering Letters**<br>
more rational methods for dynamic feature point removal,<br>
such as employing advanced selection algorithms.<br>
Additionally, we will investigate techniques to enhance the "YOLOV?: Trainable bag<br>
system' **Engineering Letters**<br>
more rational methods for dynamic feature point removal,<br>
such as employing advanced selection algorithms.<br>
Additionally, we will investigate techniques to enhance the "YOLOv7: Trainable laystem's ab **Engineering Letters**<br>
more rational methods for dynamic feature point removal,<br>
such as employing advanced selection algorithms.<br>
Additionally, we will investigate techniques to enhance the<br>
system's ability to detect a h **Engineering Letters**<br>
such as employing advanced selection algorithms.<br>
System's ability to detect a high volume of fast-moving<br>
detection network to enhance its performance.<br>
PETERENCES<br>
PETERENCES<br>
PETERENCES<br>
PETERENCE **Engineering Letters**<br>
more rational methods for dynamic feature point removal,<br>
such as employing advanced selection algorithms.<br>
Additionally, we will investigate techniques to enhance the<br>
system's ability to detect a h **Engineering Lett**<br>
more rational methods for dynamic feature point removal,<br>
such as employing advanced selection algorithms.<br>
Additionally, we will investigate techniques to enhance the<br>
system's ability to detect a high more rational methods for dynamic feature point removal,<br>
such as employing advanced selection algorithms.<br>
Additionally, we will investigate techniques to enhance the<br>
system's ability to detect a high volume of fast-movi estimational methods for dynamic feature point removal,<br>
information (ICETCI). IEE<br>
itionally, we will investigate techniques to enhance the<br>
"YOLOV?: Trainable bag<br>
entime object to the tect of the tect of the improving t Expectional methods for dynamic feature point removal.<br>
as employing advanced selection algorithms<br>
itionally, we will investigate techniques to enhance the<br>
em's ability to detect a high volume of fast-moving<br>
cts, such a such as employing advanced selection algorithms. If the method (CETCI). IE Additionally, we will investigate techniques to enhance the myOLOV? Trainable bags and the system's ability to detect a high volume of fast-moving as employing advanced selection algorithms.<br>
itionally, we will investigate techniques to enhance the<br>
em's ability to detect a high volume of fast-moving<br>
tets, such as further improving the structure of the<br>
cition netw Additionally, we will investigate techniques to enhance the<br>system's ability to detect a high volume of fast-moving<br>objects, such as further improving the structure of the<br>detection network to enhance its performance.<br>
The 29 External inertial navigation system. The Computer Vision and Computer Vision network to enhance its performance.<br>
EFERENCES<br>
Smith, Randall C., and Peter Cheeseman. "On the representation and again." Freedotics Research

### **REFERENCES**

- 
- 
- 
- cts, such as further improving the structure<br>
ction network to enhance its performance.<br>
REFERENCES<br>
Smith, Randall C., and Peter Cheeseman. "On the repres<br>
estimation of spatial uncertainty." The International<br>
Robotics R EXEREMENCES<br>
26 Liu, Wei, et al. "Ssd: S<br>
26 detection network to enhance its performance.<br>
26 Vision-ECCV 2016: 14<br>
26 Vision-ECCV 2016: 14<br>
26 Netherlands, October 11-<br>
27 Ding, \*\*aohan, et al. "Robust Schoten and estima CHOI THEVACK COM THE CONTROLL CON REFERENCES<br>
Smith, Randall C., and Peter Cheeseman. "On the representation and Publishing<br>
estimation of spatial uncertainty." The International Journal of again." Proceedings<br>
Robotics Research 5.4 (1986): 56-68.<br>
Chen, W 12462-12471. [1] Smith, Randall C., and Peter Cheeseman. "On the representation and again." Proceedings of spatial uncertainty." The International Journal of vision and Pattern Recogness Research 5.4 (1986): 56-68.<br>
[23] Chen, Weifeng Smuth, Randal C., and Peter Cheeseman. "On the representation and again." Proceedings of "<br>
Robotics Research 5.4 (1986): 56-68. <br>
Robotics Research 5.4 (1986): 56-68. <br>
Chen, Weifeng, et al. "An overview on visual slam: F estimation of spatial uncertainty." The International Journal of Vision<br>
Robotics Research 5.4 (1986): 56-68.<br>
Chen, Weifeng, et al. "An overview on visual slam: From tradition to module<br>
Seok, Hochang, and Jongwoo Lim. "R Robotics Research 5.4 (1986): 56-68.<br>
[27] Chen, Weifreig, et al. "An overview on visual slam: From tradition to module for convertions<br>
[3] Seok, Hochang, and Jongwoo Lim. "ROVINS: Robust omnidirectional<br>
[31] Seok, Hocha Chen, Weifeng, et al. "An overview on visual siam: From tradition to<br>
scock, Hochang, and Jongwoo Lim. "ROVINS: Robust omnidirectional<br>
Seok, Hochang, and Jongwoo Lim. "ROVINS: Robust omnidirectional<br>
129] Kundu, Abhijit, semantic." Remote Sensing 14.13 (2022): 3010. Con M<br>
Senok, Hochang, and Jongwoo Lim. "ROVINS: Robust omnidirectional [29] Kuno<br>
visual inertial navigation system." IEEE Robotics and Automation [29] Kun<br>
Letters 5.4 (2020) [7] Seok, Hochang, and Jongwoo Lim. "ROVINS: Robust omindirectional [29] Kundu, Abhijit, K. I<br>
visual inertial navigation system." IEEE Robotics and Automation with the United States:  $\frac{1}{2}$  Letters 5.4 (2020): 6225-623 visual inertial navigation system." IEEE Robotics and Automation<br>
Letters 5.4 (2020): 6225-6232.<br>
Letters 5.4 (2020): 6225-6232.<br>
This, Chaoyang, et al. "Robust vision-aided inertial navigation system conference on Intelli Letters 5.4 (2020): 6225-6232.<br>
Zhai, Chaoyang, et al. "Robust vision-aided inertial navigation system<br>
for protection against ego-motion uncertainty of unmanned ground<br>
vehicle." IEEE Transactions on Industrial Electronic [4] Zhai, Chaoyang, et al. "Robust vision-aided inertial navigation system Conference on Intelligent of protection against ego-motion uncertainty of umaned ground [30] Sturm, Jürgen, et al. "Monocular visual-inertial state
- 
- Li, Peiliang, et al. "Monocular visual-inertial state estim<br>mobile augmented reality." 2017 IEEE International Sympo<br>Mixed and Augmented Reality (ISMAR). IEEE, 2017.<br>Gupta, Abhishek, and Xavier Fernando. "Simultaneous loo<br>
- for protecton against ego-motion uncertainty of unmanned ground [30] Sturm, Jürgen, et vehicle." IEEE Transactions on Industrial Electronics 68.12 (2020): SLAM systems."<br>
12462-12471. <br>
Li, Peiliang, et al. "Monocular visu vehice." IEEE Transactions on Industrial Electronics 68.12 (2020):<br>
12462-12471.<br>
12462-12471. Peliang, et al. "Monocular visual-inertial state estimation for<br>
mobile augmented reality." 2017 IEEE International Symposium o 12462-12471.<br>
i. Pelilang, et al. "Monocular visual-inertial state estimation for<br>
Li, Pelilang, et al. "Monocular visual-inertial state estimation for<br>
mobile augmented Reality (ISMAR). IEEE, 2017.<br>
Gupta, Abhishek, and X mobile augmented reality." 2017 IEEE International Symposium on<br>
Mixed and Augmented Reality (ISMAR). IEEE, 2017.<br>
[6] Gupta, Abhishek, and Xavier Fernando. "Simultaneous localization<br>
and map\*\* (slam) and data fusion in u Mixed and Augmented Reality (ISMAR). IEEE, 2017.<br>
Gupta, Abhishek, and Xavier Fermando. "Simultaneous localization<br>
and map\*\* (slam) and data fusion in unmanned acrial vehicles: Recent<br>
advances and challenges." Drones 6.4 Gupta, Abhishek, and Xavier Fernando. "Simultaneous localization<br>and map\*\*\* (slam) and data faision in ummaned aerial vehicles: Recent<br>advances and challenges." Drones 6.4 (2022): 85.<br>Yin, Hesheng, et al. "Dynam-SLAM: An a and map\*\* (slam) and data fusion in unmanned aerial vehicles: Recent<br>
avances and challenges." Drones 6.4 (2022): 85.<br>
[7] Yin, Hesheng, et al. "Dynam-SLAM: An accurate, robust stereo<br>
visual-inertial SLAM method in dynam advances and challenges." Drones 6.4 (2022): 85.<br>
Yin, Hesheng, et al. "Dynam-SLAM: An accurate, robust stereo<br>
visual-inertial SLAM method in dynamic environments." IEEE<br>
Transactions on Robotics 39.1 (2022): 289-308.<br>
Ni
- 
- 
- 
- 4076-4083.<br>[12] Yuan, Xun, and Song Chen. "Sad-slam: A visual slam based on
- visual-inertial SLAM method in dynamic environments." IEEE<br>
Transactions on Robotics 39.1 (2022): 289-308.<br>
[8] Nistér, David. "Preemptive RANSAC for live structure and motion<br>
estimation." Machine Vision and Applications Transactions on Robotics 39.1 (2022): 289-308.<br>
Nistér, David. "Preemptive RANSAC for live structure and motion<br>
estimation." Machine Vision and Applications 16.5 (2005): 321-329.<br>
Campos, Carlos, et al. "Orb-slam3: An acc
- Nister, David. "Preemptive RANSAC for live structure and motion<br>estimation." Machine Vision and Applications 16.5 (2005): 321-329.<br>Campos, Carlos, et al. "Orb-slam3: An accurate open-source library for<br>visual, visual-inert estimation." Machine Vision and Applications 16.5 (2005): 321-329.<br>
[9] Campos, Carlos, et al. "Ob-slam3: An accurate open-source library for<br>
visual, visual-inertial, and multimap slam." IEEE Transactions on<br>
Robotics 37. Campos, Carlos, et al. "Orb-slam3: An accurate open-source library for<br>visual, visual-inertial, and multimap slam." IEEE Transactions on<br>Robotics 37.6 (2021): 1874-1890.<br>Yu, Chao, et al. "DS-SLAM: A semantic visual SLAM to visual, visual–inertial, and multimap slam." IEEE Transactions on<br>
Robotics 37.6 (2021): I874-1890.<br>
Yu, Chao, et al. "DS-SLAM: A semantic visual SLAM towards<br>
dynamic environments." 2018 IEEE/RSJ International Conference Robotics 37.6 (2021): 1874-1890.<br>
Yu, Chao, et al. "DS-SLAM: A semantic visual SLAM tow<br>
dynamic environments." 2018 IEEE/RSJ International Conference<br>
Intelligent Robots and Systems (IROS). IEEE, 2018.<br>
Bescos, Berta, et [10] Yu, Chao, et al. "DS-SLAM: A semantic visual SLAM towards<br>dynamic environments." 2018 IEEERSJ International Conference on<br>Intelligent Robots and Systems (IROS). IEEE, 2018.<br>[11] Bescos, Berta, et al. "DynaSLAM: Tracki dynamic environments." 2018 IEEE/RSJ International Conference on<br>Intelligent Robots and Systems (IROS). IEEE, 2018.<br>
Bescos, Berta, et al. "DynaSLAM: Tracking, map\*\*, and inpainting in<br>
dynamic scenes." IEEE Robotics and A Intelligent Robots and Systems (IROS). IEEE, 2018.<br>
[11] Bescos, Berta, et al. "DynaSLAM: Tracking, map<sup>\*\*</sup>, and inpainting in<br>
dynamic scenes." IEEE Robotics and Automation Letters 3.4 (2018):<br>
4076-4083.<br>
[12] Yuan, Xun, Bescos, Berta, et al. "DynaSLAM: Tracking, map\*\*, and inpainting in<br>dynamic scenes." IEEE Robotics and Automation Letters 3.4 (2018):<br>4076-4083.<br>Yuan, Xun, and Song Chen. "Sad-slam: A visual slam based on<br>semantic and dept dynamic scenes." IEEE Robotics and Automation Letters 3.4 (2018):<br>4076-4083.<br>Yuan, Xun, and Song Chen. "Sad-slam: A visual slam based on<br>semantic and depth information." 2020 IEEE/RSJ International<br>Conference on Intelligen 4076-4083.<br>
[12] Yuan, Xun, and Song Chen. "Sad-slam: A visual slam based on<br>
semantic and depth information." 2020 IEEE/RSJ International<br>
Conference on Intelligent Robots and Systems (IROS). IEEE, 2020.<br>
[13] Badrinaraya Yuan, Xun, and Song Chen. "Sad-slam: A visual slam based on<br>
Scenference on Intelligent Robots and Systems (IROS). IEEE, 2020.<br>
Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A<br>
deep convolutional encod semantic and depth information." 2020 IEEE/RSJ International<br>
Conference on Intelligent Robots and Systems (IROS). IEEE, 2020.<br>
[13] Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A<br>
deep convolutional Conference on Intelligent Robots and Systems (IROS). IEEE, 2020.<br>
Badrinaryanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A<br>
deep convolutional encoder-decoder architecture for image<br>
segmentation." IEEE Transact
- 
- eep convolutional encode-accooler attenties for lingual<br>segmentation." IEEE Transactions on Pattern Analysis and Machine<br>Intelligence 39.12 (2017): 2481-2495.<br>[14] He, Kaiming, et al. "Mask r-cnn." Proceedings of The IEEE<br> Segmentation. EEE Transactions on Fattern Antarysis and Machine<br>Intelligence 39.12 (2017): 2481-2495.<br>He, Kaiming, et al. "Mask r-cnn." Proceedings of The IEEE<br>International Conference on Computer Vision. 2017.<br>Kupyn, Ores [19] Redmon, Joseph, et al. "You only look once: Unified, real-time object Tre, Kamillig, et al. "Massi 1-coint" Froceedings of The EEE Conference on Computer Vision. 2017.<br>
Kupyn, Orest, et al. "Deblurgan-v2: Deblurring (orders-of-magnitude)<br>
faster and better." Proceedings of The IEEE/CVF Inter International Contentere off Computer Vision. 2017.<br>
Kupyn, Orest, et al. "Deblurgan-V2: Deblurring (orders-of-magn<br>
faster and better." Proceedings of The IEEE/CVF Interna<br>
Conference on Computer Vision. 2019.<br>
Rublee, Et
- 
- 
- (2018).<br>[18] Girshick, Ross. "Fast r-cnn." Proceedings of The IEEE International
- 
- [20] Zhong, Fangwei, et al. "Detect-SLAM: Making object detection and asset and betate. Toceecungs of The EEEPC v<sup>r</sup> mematorial Conference on Computer Vision. 2019.<br>
Rublee, Ethan, et al. "ORB: An efficient alternative to SIFT or SURF." 2011 International conference on computer vision. leee,
- Collecter of Computer Vision. 2013.<br>
Rublee, Ethan, et al. "ORB: An efficient alternative to SIFT or SURF." 2011 International conference on computer vision. Icee, 2011.<br>
Voulodimos, Athanasios, et al. "Deep learning for c [21] CONGREGITER THE COLUTE CONGREGITED TO SURF. "2011 International conference on computer vision. Ice, 2011.<br>
[17] Voulodimos, Athanasios, et al. "Deep learning for computer vision: A<br>
brief review." Computational Intell SOKET 2011 International conference on compluer vision: Leee, 2011.<br>
Voulodimos, Athanasios, et al. "Deep learning for computer vision: A<br>
brief review." Computational Intelligence and Neuroscience 2018<br>
(2018).<br>
Girshick, volucions, Anianasios, et al. Deep learning for computer vision: A<br>
brief review." Computational Intelligence and Neuroscience 2018<br>
(2018).<br>
Girshick, Ross. "Fast r-cnn." Proceedings of The IEEE International<br>
Conference 2019.<br>
[22] Kupyn, Orest, et al. "Deblurgan: Blind motion deblurring using (2018).<br>
[18] Grishick, Ross. "Fast r-cnn." Proceedings of The IEEE International<br>
[19] Redmon, Joseph, et al. "You only look once: Unified, real-time object<br>
detection." Proceedings of The IEEE Conference on Computer Visi Girshick, Ross. "Fast r-cnn." Proceedings of The IEEE International<br>Conference on Computer Vision. 2015.<br>Redmon, Joseph, et al. "You only look once: Unified, real-time object<br>detection." Proceedings of The IEEE Conference Conference on Computer Vision. 2015.<br>
Redmon, Joseph, et al. "You only look once: Unified, real-time object<br>
detection." Proceedings of The IEEE Conference on Computer Vision<br>
and Pattern Recognition. 2016.<br>
Zhong, Fangwei [19] Redmon, Joseph, et al. "You only look once: Unified, real-time object<br>dectrion." Proceedings of The IEEE Conference on Computer Vision<br>and Pattern Recognition. 2016.<br>[20] Zhong, Fangwei, et al. "Detect-SLAM: Making ob detection." Proceedings of The IEEE Conference on Computer Vision<br>and Pattern Recognition. 2016.<br>
Zhong, Fangwei, et al. "Detect-SLAM: Making object detection and<br>
SLAM mutually beneficial." 2018 IEEE Winter Conference on<br> and Pattern Recognition. 2016.<br>
Zhong, Fangwei, et al. "Detect-SLAM: Making object detection and<br>
SLAM mutually beneficial." 2018 IEEE Winter Conference on<br>
Applications of Computer Vision (WACV). IEEE, 2018.<br>
Schorghuber, [20] Zhong, Fangwei, et al. "Detect-SLAM: Making object detection and SLAM mutually beneficial." 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018.<br>
[21] Schorghuber, Matthias, et al. "SLAMA SLAM mutually beneficial." 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018.<br>Schorghuber, Matthias, et al. "SLAMANTIC-leveraging semantics to improve VSLAM in dynamic environments." Proceed
- 
- 
- 

- **Exercise Scheme Conference on Electronic Technology, Communication and Information (ICETCI). IEEE, 2021.**<br> **Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.**<br> **WOLOV7: Trainable bag-of-freebies sets new state Example 18 Example 10**<br>
Conference on Electronic Technology, Communication an<br>
Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liae<br>
"YOLOv7: Trainable bag-of-freebies sets new s **Exercise 125**<br> **Exercise Conference on Electronic Technology, Communication and Information (ICETCI). IEEE, 2021.**<br>
[25] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of-freebies se **Exercise Setters**<br>
Conference on Electronic Technology, Communication and<br>
Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of-freebies sets new stat **Exercise Conference on Electronic Technology, Communication and Information (ICETCI). IEEE, 2021.**<br>Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>"YOLOv7: Trainable bag-of-freebies sets new state-of-the-art **Example 18 Conference computer Computer**<br>
Conference on Electronic Technology, Communication and<br>
Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of
- **Exercise Shot Manuson**<br> **Exercise Conference** on Electronic Technology, Communication and<br>
Iformation (ICETCI). IEEE, 2021.<br>
[25] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of-fr **Exercise Conference Conference** on Electronic Technology, Communication and Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of-freebies sets new sta **Etters**<br>
Conference on Electronic Technology, Communication and<br>
Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of-freebies sets new state-of-the-a Conference on Electronic Technology, Communication<br>
Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark<br>
"YOLOv7: Trainable bag-of-freebies sets new state-of-the-a<br>
real-time object d Conference on Electronic Technology, Communication and<br>
Information (ICETCI). IEEE, 2021.<br>
[25] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOV7: Trainable bag-of-freebies sets new state-of-the-art fo Conference on Electronic Technology, Communication and Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for<br>
rea Conference on Electronic Technology, Communication and<br>Information (ICETCI). IEEE, 2021.<br>Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>"YOLOv7: Trainable bag-of-frecbies sets new state-of-the-art for<br>real-t Conference on Electronic Technology, Communication and<br>
Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for<br>
re Information (ICETCI). IEEE, 2021.<br>
Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.<br>
"YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for<br>
real-time object detectors." Proceedings of The IEEE/CVF C Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao<br>"YOLOv7: Trainable bag-of-freebies sets new state-of-the-art fo<br>real-time object detectors." Proceedings of The IEEE/CVF Conference<br>on Computer Vision and Patter "YOLOv7: Trannable bag-of-freebies sets new state-of-the-art for<br>read time object detectors." Proceedings of The IEEE/CVF Conference<br>on Computer Vision and Pattern Recognition. 2023.<br>[26] Liu, Wei, et al. "Ssd: Single shot real-time object detectors." Proceedings of The IEEE/CVF Conference<br>
on Computer Vision and Pattern Recognition. 2023.<br>
Liu, Wei, et al. "Ssd: Single shot multibox detector." Computer<br>
Vision-ECCV 2016: 14th European Confe on Computer Vision and Pattern Recognition. 2023.<br>
Liu, Wei, et al. "Ssd: Single shot multibox detector." Computer<br>
Vision-ECCV 2016: 14th European Conference, Amsterdam, The<br>
Netherlands, October 11–14, 2016, Proceedings,
- 
- 
- Liu, Wei, et al. "Ssd: Single shot multibox detector." Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016. Vision-ECCV 2016: 14th European Conterence, Amsterdam, The<br>Netherlands, October 11-14, 2016, Proceedings, Part I 14. Springer<br>International Publishing, 2016.<br>The International Publishing, 2016.<br>The IEEE/CVF Conference on C Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016.<br>Ding, \*\*aohan, et al. "Repvyg: Making vgg-style convnets great<br>again." Proceedings of The IEEE/CVF Conference on Computer<br>V International Publishing, 2016.<br>
Ding, \*\*aohan, et al. "Repygg: Making vgg-style convnets great<br>
again." Proceedings of The IEEE/CVF Conference on Computer<br>
Vision and Pattern Recognition. 2021.<br>
Yang, Lingxiao, et al. "Si
-