

Article

Molecular-Clump Detection Based on an Improved YOLOv5 Joint Density Peak Clustering

Jin-Bo Hu ^{1,2}, Yao Huang ^{1,3,*}, Sheng Zheng ^{1,3,*}, Zhi-Wei Chen ⁴ , Xiang-Yun Zeng ^{1,3}  and Xiao-Yu Luo ^{1,3} and Chen Long ^{1,3} 

¹ Center for Astronomy and Space Sciences, China Three Gorges University, Yichang 443000, China; jinbohu@ctgu.edu.cn (J.-B.H.); xyzeng2018@ctgu.edu.cn (X.-Y.Z.); lxy@ctgu.edu.cn (X.-Y.L.); longc25@ctgu.edu.cn (C.L.)

² College of Computer and Information Technology, China Three Gorges University, Yichang 443000, China

³ College of Science, China Three Gorges University, Yichang 443000, China

⁴ Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, China; zwchen@pmo.ac.cn

* Correspondence: huangyao@ctgu.edu.cn (Y.H.); zsh@ctgu.edu.cn (S.Z.)

Abstract: The detection and analysis of molecular clumps can lead to a better understanding of star formation in the Milky Way. Herein, we present a molecular-clump-detection method based on improved YOLOv5 joint Density Peak Clustering (DPC). The method employs a two-dimensional (2D) detection and three-dimensional (3D) stitching strategy to accomplish the molecular-clump detection. In the first stage, an improved YOLOv5 is used to detect the positions of molecular clumps on the Galactic plane, obtaining their spatial information. In the second stage, the DPC algorithm is used to combine the detection results in the velocity direction. In the end, the clump candidates are positioned in the 3D position-position-velocity (PPV) space. Experiments show that the method can achieve a high recall of 98.41% in simulated data made up of Gaussian clumps added to observational data. The efficiency of the strategy has also been demonstrated in experiments utilizing observational data from the Milky Way Imaging Scroll Painting (MWISP) project.

Keywords: YOLOv5; density peak clustering; molecular clouds; clump detection



Citation: Hu, J.-B.; Huang, Y.; Zheng, S.; Chen, Z.-W.; Zeng, X.-Y.; Luo, X.-Y.; Long, C. Molecular-Clump Detection Based on an Improved YOLOv5 Joint Density Peak Clustering. *Universe* **2023**, *9*, 480. <https://doi.org/10.3390/universe9110480>

Academic Editor: Jason McEwen

Received: 29 September 2023

Revised: 3 November 2023

Accepted: 7 November 2023

Published: 11 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As one of the most essential components of the interstellar medium, molecular clouds consist primarily of molecular gas mixed with small amounts of atoms, ions, and dust [1]. The scale of molecular clouds spans from giant clouds with tens of pc to dense clumps with less than 0.1 pc [2]. Molecular clouds usually present complex and hierarchical structures. Clumpiness is a universal property of molecular clouds and plays a key role in understanding the fragmentation of cloud complexes into sub-clouds finally small enough to form individual stars [3]. Modern astronomy recognizes dense clumps in the molecular clouds as the key sites of star formation [4]. Thus, the detection and analysis of dense clumps in molecular clouds can better elucidate star formation and matter cycling in the Milky Way [5].

Several researchers have already studied automatic detection algorithms for molecular clumps. In 1990, Stutzki and Guesten (1990) [3] proposed the GaussClumps algorithm and applied it to the M17 molecular cloud to detect 170 clumps. In 1994, Williams et al. (1994) [6] proposed the ClumpFind algorithm and tested it on RMC and MMC. They detected 83 clumps in the RMC and 78 clumps in the two observed regions of MMC. In 2015, Berry (2015) [7] proposed the FellWalker algorithm, which has been utilized by Kirk et al. (2016) [8] to detect a total of 915 clumps in three subregions of the Orion B molecular cloud and obtain basic physical information, such as the fluxes and sizes of these clumps. In 2022, Luo et al. (2022) [9] proposed the LDC algorithm, which has detected 658 clumps in the M16 molecular cloud. Astronomers need to set the initial parameters of these detection algorithms according to the observational instrument and the morphological features of the

target during detection. The detection results are closely related to the parameter settings. Before detection, astronomers need to repeatedly adjust the parameters to optimize the algorithm or manually verify the detection results and thus ensure accuracy. To accomplish the task of large-scale molecular clump census from the rapidly growing CO molecular-cloud observation, the research, and development of detection algorithms with higher efficiency are needed.

With the rapid development of artificial intelligence technology, deep-learning methods represented by convolutional neural networks (CNNs) have extensively been used in image and machine vision [10–15]. Deep learning also provides a viable method for extracting essential features of astronomical data. Kim and Brunner (2017) [16] proposed a CNN to classify stars and galaxies. Gonzalez et al. (2018) [17] implemented the task of galaxy detection and classification by using the YOLO target-detection algorithm. Leung and Bovy (2019) [18] designed a Bayesian neural network to measure multi-element abundances in stellar spectra. Xie et al. (2021) [19] applied a convolutional neural model to stellar spectra for detecting stars with low metal content. He et al. (2021) [20] proposed a target-detection network based on YOLOv4 to detect sources in SDSS images and a classification model called APSCnet to classify sources. Yi et al. (2022) [21] proposed an automatic detection model based on a deep-learning approach for automatically detecting low-surface-brightness galaxies from SDSS images. Cao et al. (2023) [22] used an improved Faster R-CNN framework based on deep learning to detect L-dwarf from SDSS images automatically. These new techniques enable the automatic analysis of astronomical data and demonstrate the feasibility of supervised deep learning for astronomical data mining. Therefore, an automatic clump-detection algorithm that extracts the desired types of targets directly from the observational data through supervised feature learning from clump morphological properties can be similarly designed for CO survey data.

The Milky Way Imaging Scroll Painting (MWISP) project [23] is a large-scale CO survey project carried out by the Purple Mountain Observatory of the Chinese Academy of Sciences by using the DLH-13.7 m millimeter-wave radio telescope, which simultaneously observes ^{12}CO , ^{13}CO , and $\text{C}^{18}\text{O}(1-0)$ lines emission. The first phase of MWISP has acquired a large number of molecular-cloud observational data, providing sufficient training samples for the deep-learning-based clump detection algorithm. This paper proposes a clump-detection method based on an improved YOLOv5 joint Density Peak Clustering (DPC) [24]. The method initially locates the clumps in the position-position (PP) coordinate representing the Galactic longitude and latitude by using the improved YOLOv5. Subsequently, DPC is used to cluster the clumps in the velocity direction, thereby ultimately realizing the clump detection in position-position-velocity (PPV) three-dimensional (3D) space. This method utilizes supervised deep learning to detect clumps by labeling the areas of interest. The detection results are directly related to the labeling scheme and feature learning of the data. During detection, the 3D information where the target candidate is located can be obtained quickly with far fewer parameters tuning during detection.

The paper is organized as follows. Section 2 describes the generation of the experimental data. Section 3 details the theory of the clump-detection algorithm based on the improved YOLOv5 joint DPC. Section 4 presents the training process and the main experimental results in PPV space. Finally, conclusions are drawn in Section 5.

2. Data

To better train the improved YOLOv5 and comprehensively evaluate the performance of the proposed method, a large number of clumps are required. Synthesized data are designed for the quantitative evaluation of detection performance. Synthesized data are composed by randomly adding simulated clumps to the observational data background. The simulated clumps are generated by the 3D Gaussian model, and the background is the observational $^{13}\text{CO}(1-0)$ line emission data obtained by MWISP. The simulated clumps, the observational data, and the synthesized data are described separately below.

2.1. Simulated Clumps

Localized high densities characterize the clumps, and Gaussian functions can be used to describe their intensity distributions [3]. Simulated clumps can be generated utilizing the 3D Gaussian model with several parameters [25], as shown in Table 1. The 3D Gaussian model is described as:

$$f(x, y, v) = A \exp \left\{ - \left[\left(\frac{\cos^2 \theta}{2\sigma_x^2} + \frac{\sin^2 \theta}{2\sigma_y^2} \right) (x - x_0)^2 + \left(\frac{\cos^2 \theta}{2\sigma_y^2} - \frac{\sin^2 \theta}{2\sigma_x^2} \right) (x - x_0)(y - y_0) + \left(\frac{\sin^2 \theta}{2\sigma_y^2} + \frac{\cos^2 \theta}{2\sigma_x^2} \right) (y - y_0)^2 + \frac{(v - v_0)^2}{2\sigma_v^2} \right] \right\} \tag{1}$$

where A represents the peak intensity of the clump, $\sigma_x, \sigma_y, \sigma_v$ represent the standard deviations on the Galactic longitude, the Galactic latitude, and the velocity, respectively, (x_0, y_0, v_0) represents the position of the center of mass of the clump, and θ represents the rotation angle on the Galactic plane. To enable generated simulated clumps with the parameter-distribution characteristics of the observational data, the statistics of the clump parameters detected in the $^{13}\text{CO}(1-0)$ line emission data of the M16 region obtained by MWISP are regarded as the input into the 3D Gaussian model. The $^{13}\text{CO}(1-0)$ line emission of M16 region ranges $15^\circ 15' < l < 18^\circ 15', 0^\circ < b < 1^\circ 30'$. LDC [9] algorithm was applied to detect the $^{13}\text{CO}(1-0)$ line emission of M16 and a total of 658 clumps have been detected. We counted the morphological parameters of these clumps and obtained the parameters range in Table 1. Five thousand simulated clumps were randomly generated, and fluxes were calculated according to the parameters in Table 1. The flux distribution is shown in Figure 1. These simulated clumps maintain consistency with the MWISP observational data and provide a more realistic representation of the detection performance.

Table 1. Parameters of the 3D Gaussian model.

Parameter Name	Explanation	Range
Peak	Peak intensity of the clump	[0.7, 15]
σ_x	Standard deviation on the Galactic longitude	$[1, 4] \times 2.3548$
σ_y	Standard deviation on the Galactic latitude	$[1, 4] \times 2.3548$
σ_v	Standard deviation on the velocity direction	$[1, 7] \times 2.3548$
(x_0, y_0, v_0)	Position of the center of mass of the clump	Randomization
θ	Rotation angle on the Galactic plane	$0^\circ - 180^\circ$

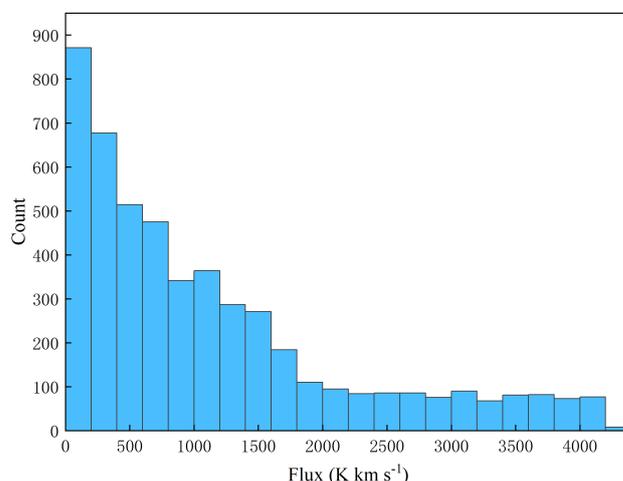


Figure 1. Distribution of the number of fluxes of 5000 randomly generated simulated clumps.

2.2. Observational Data

Three typical regions within the first, second, and third Galactic quadrants of the $^{13}\text{CO}(1-0)$ line emission data obtained from MWISP are selected as the background for generating the synthesized data. The typical noise level of $^{13}\text{CO}(1-0)$ line emission is about 0.23 K with a channel width of 0.167 km s^{-1} . Three regions all contain information in the three dimensions of Galactic longitude, Galactic latitude, and velocity. Due to the structure of the spiral arm of the Milky Way, different quadrants contain different spiral arms and have different gas distributions. The range of the Galactic plane is $3^\circ \times 2^\circ$ and the velocity range is 70 km s^{-1} , of all three backgrounds, corresponding with the size of $361 \times 241 \times 424$ pixels. The regions selected in the first Galactic quadrant range $13^\circ < l < 16^\circ$ and $-1^\circ 30' < b < 30'$, and the velocity ranges $0 \text{ km s}^{-1} < v_s. < 70 \text{ km s}^{-1}$. It belongs to the inner Milky Way area, in the direction of the galactic center. The regions selected in the second Galactic quadrant range $101^\circ < l < 104^\circ$ and $2^\circ < b < 4^\circ$, and the velocity ranges $-60 \text{ km s}^{-1} < v_s. < 10 \text{ km s}^{-1}$. The regions selected in the third Galactic quadrant range $184^\circ 30' < l < 187^\circ 30'$ and $-1^\circ < b < 1^\circ$, and the velocity ranges $-10 \text{ km s}^{-1} < v_s. < 60 \text{ km s}^{-1}$. They both belong to the outer Milky Way area. The density of the three regions is different. The background gas is dense in the first Galactic quadrant region and sparse in the third Galactic quadrant region. In the second Galactic quadrant region, the density of the background is between the first and the third Galactic quadrants. Different gas densities can reflect the detection performance of the detection algorithm in different signal-to-noise ratio environments. Figure 2 shows the velocity-integrated intensity maps for each selected background region.

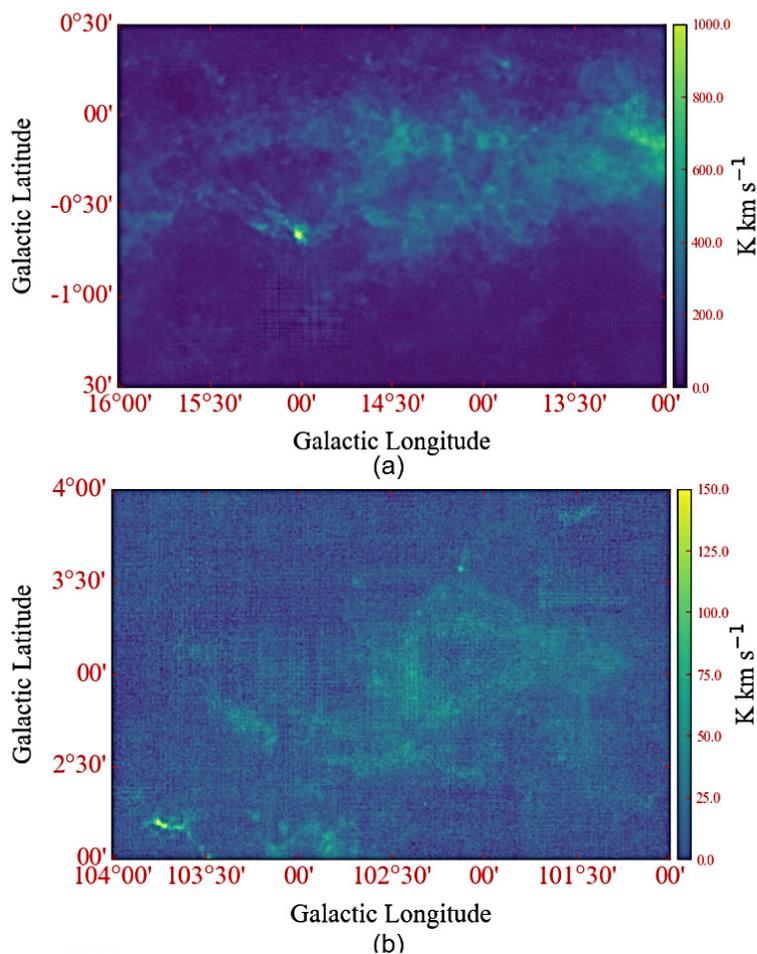


Figure 2. Cont.

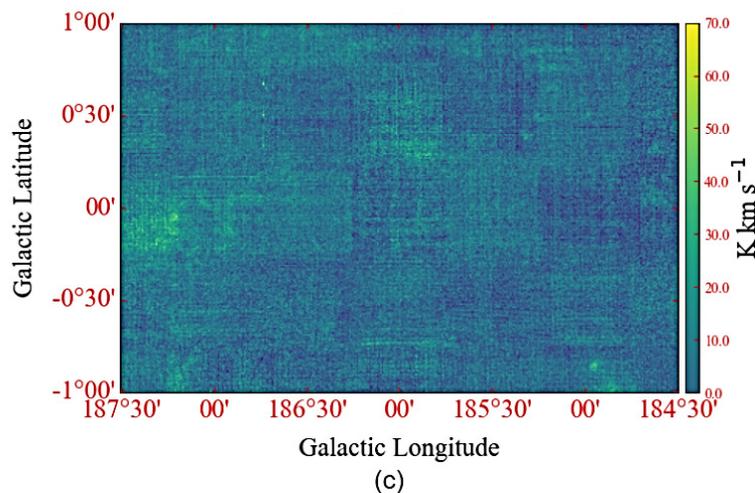


Figure 2. Velocity–integrated intensity maps for selected background regions. (a) the first Galactic quadrant with $13^\circ < l < 16^\circ$ and $-1^\circ 30' < b < 30'$; the velocity ranges $0 \text{ km s}^{-1} < v_s. < 70 \text{ km s}^{-1}$. (b) the second Galactic quadrant with $101^\circ < l < 104^\circ$ and $2^\circ < b < 4^\circ$; the velocity ranges $-60 \text{ km s}^{-1} < v_s. < 10 \text{ km s}^{-1}$. (c) the third Galactic quadrant with $184^\circ 30' < l < 187^\circ 30'$ and $-1^\circ < b < 1^\circ$; the velocity ranges $-10 \text{ km s}^{-1} < v_s. < 60 \text{ km s}^{-1}$.

2.3. Synthesized Data

One hundred simulated clumps obtained using the 3D Gaussian model are added randomly to the designated region mentioned in Section 2.2 to generate synthesized data. The added simulated clumps are uniformly distributed throughout all regions. This process is repeated three times to create the first, second, and third Galactic quadrants’ synthesized data. To ensure that each synthesized data item has 100 complete simulated clumps, it should avoid boundary regions and clump fusion during the addition process. Velocity-integrated intensity maps for three Galactic quadrants’ synthesized data are shown in Figure 3. Synthesized data also contains information in the three dimensions of Galactic longitude, Galactic latitude, and velocity. The size of synthesized data is the same as its corresponding background region size. The range of the Galactic plane is $3^\circ \times 2^\circ$ and the velocity range is 70 km s^{-1} , of all synthesized data, corresponding with the size of $361 \times 241 \times 424$ pixels.

As the synthesized data are generated, information about each added simulated clump is recorded in a clump parameter table for each synthesized data. The information contained in the table is shown in Table 2, where ID denotes the clump number. Peak1, Peak2, and Peak3 indicate the peak intensity coordinates. Cen1, Cen2, and Cen3 indicate the center of mass coordinates. Size1, Size2, and Size3 denote the axial lengths. θ represents the rotation angle on the Galactic plane. Sum denotes the total flux. Peak indicates the peak intensity.

Once the detection is complete, the results are compared with the clump parameter tables. The algorithm performance can then be quantitatively evaluated.

Table 2. Parameter information table of simulated clumps.

Parameter Name	Explanation
ID	Clump number
Peak1, Peak2, Peak3	Peak coordinates of clumps
Cen1, Cen2, Cen3	Coordinates of the center of mass of clumps
Size1, Size2, Size3	Axis lengths of clumps in the Galactic plane, and velocity direction
θ	Rotation angles of clumps on the Galactic plane
Sum	Total flux of clumps
Peak	Peak intensity of clumps

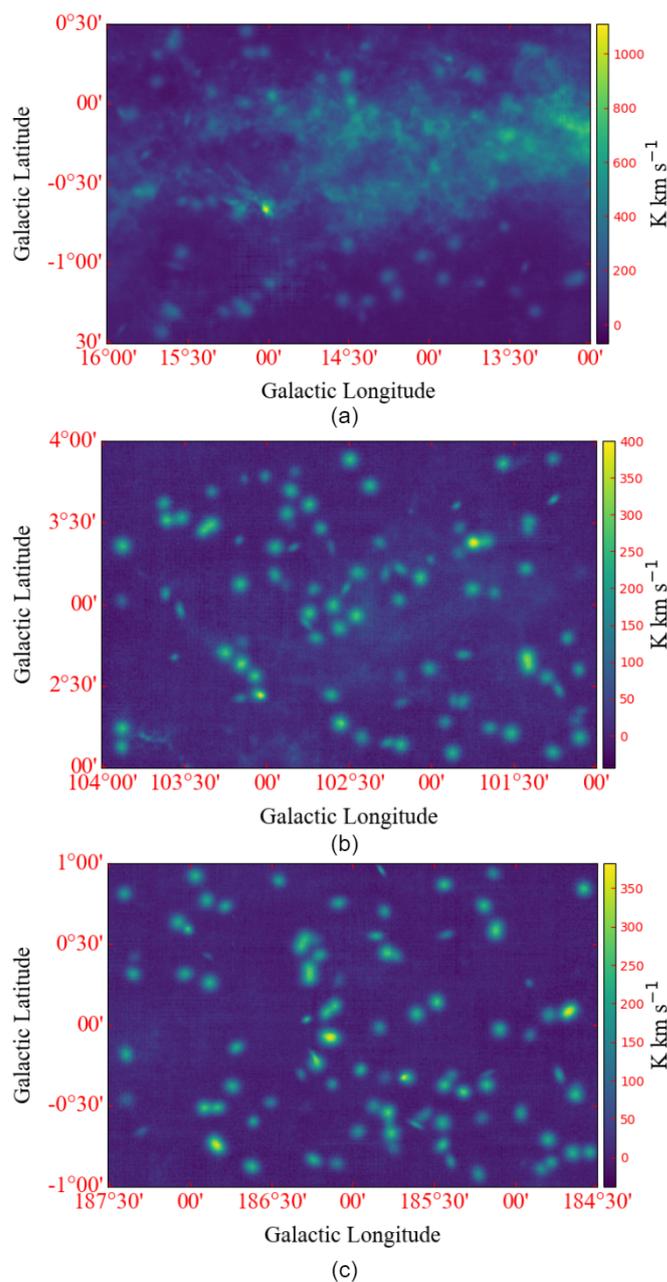


Figure 3. Velocity–integrated intensity maps of the synthesized data: (a) the first Galactic quadrant synthesized data. (b) the second Galactic quadrant synthesized data. (c) the third Galactic quadrant synthesized data.

3. Method

This paper presents a new method based on an improved YOLOv5 joint DPC for detecting molecular clumps in PPV data. Detection is divided into two stages. In the first stage, the improved YOLOv5 is used to detect the clumps in two-dimensional (2D) space and obtain their position on the Galactic plane. In the second stage, DPC is used to combine the detection results obtained in the first stage in the velocity direction, thereby helping achieve the final clump detection in 3D space. The code has been shared on Github ¹.

3.1. The Improved YOLOv5 – Molecular Clump Detection (MCD)-YOLOv5

YOLOv5 is an extensively used target-detection tool that combines speed and accuracy [26–29]. One of its key advantages is that it is a one-stage detection algorithm. It

provides faster inference and detection speeds, making it useful for handling large amounts of molecular cloud data.

YOLOv5 consists of a backbone, neck, and YOLO head. The backbone can extract the information of the input. The neck can extract features from the backbone to improve the performance of the network. The YOLO head outputs the results of the network predictions. Although YOLOv5 is a strong performer in many target-detection tasks, it has limitations in detecting small targets which are often low in pixels, small in percentage, easy to overlap, and difficult to distinguish. The molecular-cloud data have clumps with weak intensity which may be obscured by the molecular-cloud background. Owing to the rotation angle, some clumps also appear small on the Galactic plane. To enhance the ability of YOLOv5 to detect these small clumps, we introduce the Coordinate Attention (CA) module and Normalized Wasserstein Distance (NWD) loss function to YOLOv5. This improved version, called Molecular clump detection (MCD)-YOLOv5, and its architecture is shown in Figure 4.

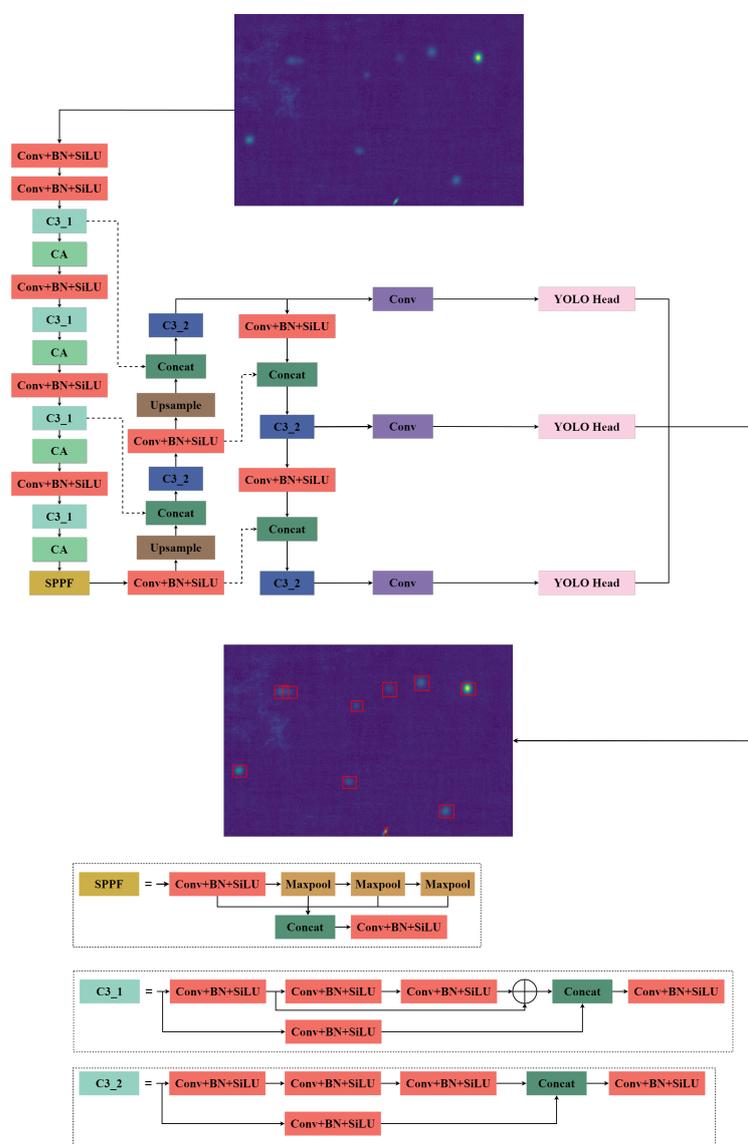


Figure 4. Architecture of MCD-YOLOv5. MCD-YOLOv5 is mainly composed of three parts: backbone, neck, and YOLO head. The backbone consists of CBS, C3_1, and CA modules. CBS is a composite convolution module, which encapsulates a convolutional layer, a batch normalization layer, and the SiLU activation function. MCD-YOLOv5 contains two types of C3 modules; C3_1 is applied in the backbone and C3_2 is applied in the neck. The C3_1 module and C3_2 module both consist of CBS.

In C3_1, the input feature map passes through three branches, while in C3_2, it only passes through two branches. The branches are finally spliced by channel and then output through a CBS module. We add a CA module after each layer of the C3_1. The neck consists of SPPF(Spatial Pyramid Pooling-Fast), and CSP-PAN(CSP-Path Aggregation Network) modules. SPPF is a spatial pyramid pooling module. SPPF encapsulates a CBS module and three maximum pooling layers. Three pooling results with input feature maps are spliced by channel and passed through a CBS module. CSP-PAN is composed of CBS and C3_2, and the feature fusion of different feature layers is realized through upsampling and downsampling, which solves the target multi-scale problem to a certain extent. The main part of the YOLO head is three detectors, that is, using mesh-based anchors to detect objects on feature maps at different scales.

3.1.1. Coordinate Attention

To improve YOLOv5’s ability to recognize the clumps with weak intensity in slices, a CA module [30] is utilized. The architecture of the CA module is shown in Figure 5. CA is a lightweight attention module, which encodes each channel of the feature map along the X and Y directions. The resulting feature maps are combined and transformed via convolutional transform to create intermediate feature maps. These feature maps are then divided into separate tensors, which are again transformed and expanded to become the value of the attention weight assignment. The CA module is added after each layer of the C3_1 module in YOLOv5, thereby increasing the number of layers from 10 to 14. This action enables a more accurate identification of regions of interest during the feature extraction of clumps.

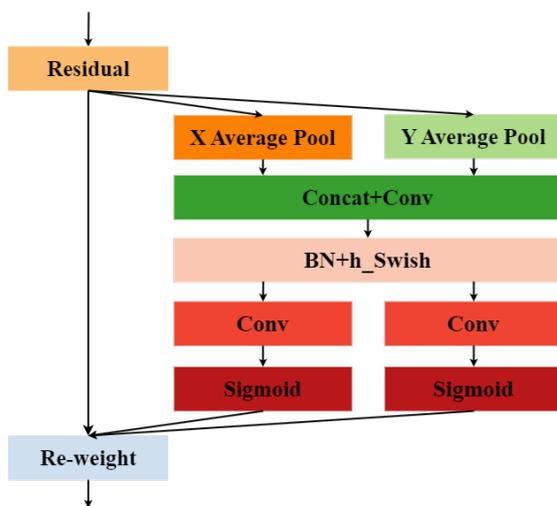


Figure 5. Architecture of Coordinate Attention module. The module is made of two average pooling layers, a convolution layer with concat operation, a batch normalization layer with an h_swish activation function, two convolution layers, and two sigmoid activation functions. The average pooling layers encode each channel of the feature map along the X and Y directions. The resulting feature maps are combined and transformed via the convolution layer with concat operation and the batch normalization layer with an h_swish activation function to create intermediate feature maps. These feature maps are then divided into separate tensors, which are again transformed and expanded by the convolution layer and sigmoid activation function to become the value of the attention weight assignment.

3.1.2. Normalized Wasserstein Distance

Detecting small clumps is challenging owing to limited appearance information, conferring difficulty in recognizing features. The use of Intersection over Union (IOU) [31] as a metric in YOLOv5 to measure the goodness of the generated anchor frames during detection also results in the poor detection of small targets. Wang et al. (2022) [32] analyzed the sensitivity of IoU to location deviations of tiny objects, and proposed NWD as a better metric for measuring the similarity between two bounding boxes. To improve the detection

performance of YOLOv5 for small target clumps, the NWD loss function is introduced. The expressions for NWD and the loss function based on NWD are presented as follows:

$$W_2^2(N_p, N_g) = \left\| \left(\left[cx_p, cy_p, \frac{w_p}{2}, \frac{h_p}{2} \right]^T, \left[cx_g, cy_g, \frac{w_g}{2}, \frac{h_g}{2} \right]^T \right) \right\|_2^2 \tag{2}$$

$$NWD(N_p, N_g) = \exp \left(-\frac{\sqrt{W_2^2(N_p, N_g)}}{C} \right) \tag{3}$$

$$L_{NWD} = 1 - NWD(N_p, N_g) \tag{4}$$

where N_p is the Gaussian distribution model of the prediction frame P; N_g is the Gaussian distribution model of the gt frame G; (cx_p, cy_p, w_p, h_p) is the coordinates of the upper left corner of the prediction frame P and the width and height information; (cx_g, cy_g, w_g, h_g) represents the coordinates of the upper left corner of the gt frame G and the width and size information, respectively; and $W_2^2(N_p, N_g)$ is the second-order Wasserstein distance between N_p and N_g . C is a constant. This paper uses the original Complete IoU Loss (CIoU) [33] of YOLOv5 and NWD loss in the molecular-clump-detection task. The weight ratio of CIoU loss and NWD loss is 7:3.

3.2. Density Peak Clustering Algorithm

The DPC algorithm was proposed by Rodriguez & Laio (2014) [24] in 2014. DPC assumes that points with low local density surround the cluster centers, which are relatively distant from others. This feature is very similar to clumps embedded in the molecular cloud with lower average density and localized density enhancement. Accordingly, this paper uses DPC to complete the splicing of clumps in the velocity direction.

DPC is implemented by dividing the data points within a specific range into a region followed by calculating the density ρ_i and distance δ_i of each data point within that region:

$$\delta_i = \begin{cases} \min_{j: \rho_j \geq \rho_i} d_{ij}, & \text{if } \rho_j \geq \rho_i \\ \max_{j=1,2,\dots,n} d_{ij}, & \text{otherwise} \end{cases} \tag{5}$$

d_{ij} denotes the Euclidean distance between data points i and j . When a point with a higher density than the current data point exists, the distance of this point is the minimum of all data points denser than the point. After setting the thresholds for density and distance, clustering centers can be determined from these data points. Clustering can also be accomplished by assigning points not clustering centers to the nearest clustering centers based on the distances of other data points from the clustering centers.

After MCD-YOLOv5, the clumps are extracted in the slices by the coordinate position information obtained from the detection, and the center of mass of the extracted region is calculated by the adaptive threshold segmentation [34]. Through the position coordinates (x_i, y_i) of the center of mass and its velocity channel information v_i , the intensity I_i of the corresponding position in the synthesized data is also extracted. These pieces of information are used in this paper to represent the density of each input data point with the following expression:

$$\rho_i = I_i \tag{6}$$

The inputs and outputs of DPC are listed in Table 3. DPC clusters the centers with position and density output by MCD-YOLOv5 and divides them into different categories. The final number of categories is the number of clumps detected.

Table 3. Input and output data for DPC.

Description	Parameters Name	Explanation
Input	x_i	Galactic longitude coordinates of the center of mass in the region detected by MCD-YOLOv5
	y_i	Galactic latitude coordinates of the center of mass in the region detected by MCD-YOLOv5
	v_i	Channels in the velocity direction detected by MCD-YOLOv5
Output	I_i	Intensity of (x_i, y_i, v_i) in the synthesized data
	numClust	Number of clustered clumps categories

3.3. MCD-YOLOv5 Joint DPC

The flowchart of the clump-detection method based on MCD-YOLOv5 joint DPC is shown in Figure 6. The molecular cloud data are sliced in the velocity direction, resulting in 2D data along the Galactic plane. These data are inputted into MCD-YOLOv5 for detection. Once all slices have been detected, MCD-YOLOv5 outputs the coordinate position information of the detected clumps. The center of mass of each detected clump can be calculated based on MCD-YOLOv5 output information. Using the center of mass and the velocity channel size where the corresponding slice is located, the intensity value of the related part on the molecular cloud data is obtained as the input to DPC. DPC clusters the centers of mass and classifies them into different categories. The region comprising clumps corresponding with center of mass points in the same category is classified as a clump. The spatial coordinates of the slices of molecular clouds belonging to the same clump in the direction of the Galactic plane and the span in the direction of velocity are statistically calculated to transform the PP data into PPV data. The output of MCD-YOLOv5 joint DPC is 3D cubic regions containing clumps. By obtaining information about each 3D region, we can calculate the centroid (Cen_i) and size ($Size_i$) of the molecular cloud clumps on different axes i ($i = 1, 2, 3$). The definition of the centroid is as follows:

$$Cen_i = \frac{\sum_{j=1}^{n_k} (I_j \cdot x_j)}{\sum_{j=1}^{n_k} I_j} \tag{7}$$

the $Size_i$ of the clump on axis i is defined as:

$$Size_i = \sqrt{\frac{\sum_{j=1}^{n_k} (I_j \cdot x_j^2)}{\sum_{j=1}^{n_k} I_j} - \left(\frac{\sum_{j=1}^{n_k} (I_j \cdot x_j)}{\sum_{j=1}^{n_k} I_j}\right)^2} \tag{8}$$

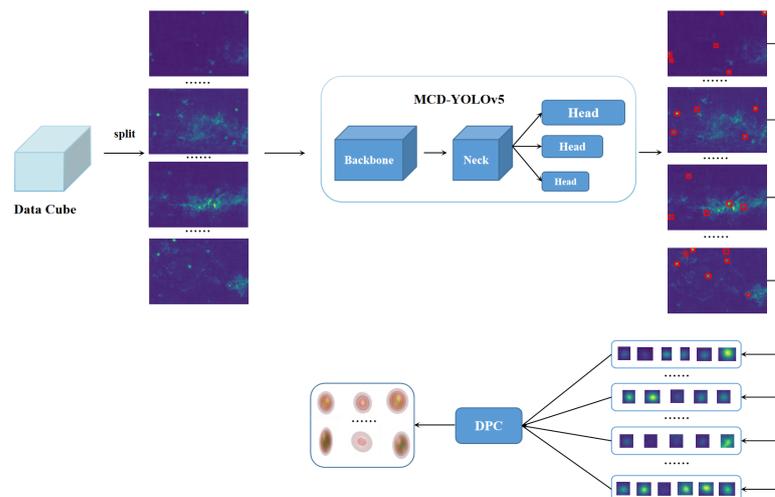


Figure 6. Flowchart of the clump-detection method based on MCD-YOLOv5 joint DPC. The four branches represent any four slices in a given data cube. Each slice achieves 2D detection by MCD-YOLOv5 to obtain the anchor information of the target on the corresponding slice. The rest slices are indicated by ellipses. Heads in different size fonts indicate different scales of detection heads.

4. Experiments and Discussion

To verify the effectiveness and reliability of the detection method based on MCD-YOLOv5 joint DPC, MCD-YOLOv5 is trained by using synthesized data from the first and third Galactic quadrants. Its performance in detecting clumps is tested using the synthesized data from the second Galactic quadrant. The training process of MCD-YOLOv5, the effectiveness of DPC, and the test results are presented in the following sections.

4.1. Evaluation of Indicators

The performance of MCD-YOLOv5 for detecting clumps on the 2D Galactic plane is quantified by using precision, recall, and average precision (AP), which are the most commonly used in target-detection tasks. Using the recall rate to evaluate the 3D detection performance of the clump-detection method based on MCD-YOLOv5 joint DPC. The precision, recall, and AP expressions are as follows:

$$precision = \frac{TP}{TP + FP} \quad (9)$$

$$recall = \frac{TP}{TP + FN} \quad (10)$$

$$AP = \int_0^1 P(r) \quad (11)$$

In calculating precision and recall, true positive (TP) denotes the number of clumps predicted to be clumps that are actually clumps, false positive (FP) denotes the number of clumps predicted to be clumps but are not actually clumps, and false negative (FN) denotes the number of clumps predicted to be not clumps that are actually clumps. In calculating AP, $P(r)$ denotes the maximum precision value when the recall takes the corresponding value.

4.2. MCD-YOLOv5 Training and DPC

4.2.1. MCD-YOLOv5 Dataset Generation

The slices generated by intercepting along the velocity channel of the synthesized data are used as samples in the dataset. Each sample has only one velocity channel and contains information in two dimensions, Galactic longitude, and Galactic latitude, with a pixel size of 361×241 . A total of 45 first-Galactic-quadrant and 45 third-Galactic-quadrant synthesized data cubes with 9000 simulated clumps to generate the dataset used for the training and validation of MCD-YOLOv5. There are a total of 37,350 samples in this dataset. The samples obtained from the first-Galactic-quadrant and third-Galactic-quadrant synthesized data are shown in Figure 7a and Figure 7b, respectively.

Considering that MCD-YOLOv5 is a supervised deep-learning method, the location information on the clumps in each sample needs to be provided. The position information of clumps in the training samples can be obtained by the clump parameter table mentioned in Section 2.3. The specific process for obtaining the label information is as follows. The center of mass (Cen1 and Cen2) of the simulated clumps on the Galactic plane is taken as the starting point. To form a rectangle that can cover the clumps, the pixel distance with Size1 is expanded upward and downward, and the pixel distance with Size2 is expanded to the left and the right. Recording the coordinate position information of the rectangles gives information about the labeling of clumps on a sample. We can obtain the position information of the same simulated clump appearing in the sample by taking the center of mass coordinate Cen3 of the simulated clump in the velocity direction as the midpoint and expanding the width of Size3 forward and backward. Figure 7 shows that the simulated clumps are accurately labeled by the above method. Figure 7c shows the labeling of the synthesized data slices in the third Galactic quadrant. Figure 7d shows the labeling of the synthesized data slices in the first Galactic quadrant.

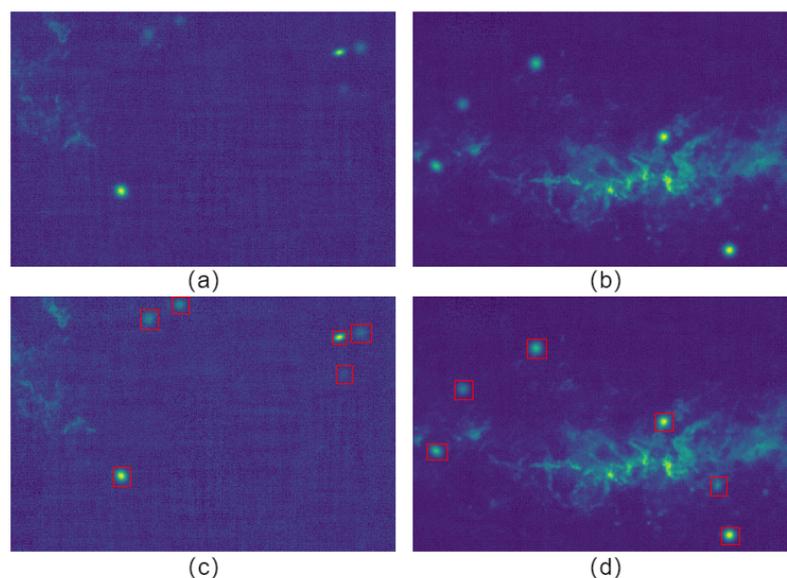


Figure 7. Examples from the MCD-YOLOv5 dataset and the annotation information: (a) synthesized data in the third Galactic quadrant, (b) synthesized data in the first Galactic quadrant, (c) annotation information on (a), and (d) annotation information on (b). The red-labeled boxes represent the clumps.

4.2.2. Training MCD-YOLOv5

To train MCD-YOLOv5 effectively, the dataset is divided into a training set and a test set at a ratio of 8:2, and takes 10% from the training set as the validation set. The training set is used to train MCD-YOLOv5, the validation set is used to verify the performance metrics of MCD-YOLOv5 during training, and the test set is used to test the detection effect of MCD-YOLOv5 after training is completed. During training, the number of training rounds is set to 300, the amount of batch training data is 128, the training momentum is 0.9, the initial learning rate is 0.001, and the weight decay is 0.0005. The stochastic gradient descent (SGD) serves the optimization function to train MCD-YOLOv5. The default parameters of YOLOv5 were used for training in this paper.

To verify the improved performance of MCD-YOLOv5 compared with YOLOv5 in detecting clumps, we train MCD-YOLOv5 and YOLOv5 on the same dataset. Figure 8 shows the variation in localization loss and confidence loss. The curve variation shows that MCD-YOLOv5 and YOLOv5 are not overfitted or underfitted, and MCD-YOLOv5 decreases more smoothly during training. The variation curves of the precision and recall on the validation set with the number of training rounds are shown in Figure 9a,b. Figure 9c demonstrates the variation of AP in MCD-YOLOv5 and YOLOv5 during the training when IOU is set to 0.5. IOU denotes the intersection and concurrency ratio of the area of the predicted frame to the area of the actual frame. In general, the predicted frames produced by the model are recognized to be correct only if the IOU is greater than or equal to the set threshold. The curve changes show that the precision of MCD-YOLOv5 does not considerably differ from that of YOLOv5, but its recall rate exceeds that of YOLOv5.

We select a slice containing smaller clumps and input it into the trained MCD-YOLOv5 and YOLOv5 at the same time for detection. Results are shown in Figure 10. By comparison, MCD-YOLOv5 detects all seven clumps in the slices, whereas YOLOv5 misses a smaller clump labeled 2. This experiment illustrates the improved small target detection performance of the MCD-YOLOv5. Table 4 shows the precision, recall, and AP of MCD-YOLOv5 and YOLOv5 in the test set. By comparison, MCD-YOLOv5 performs better than YOLOv5 in detecting clumps when processing the same molecular cloud data slices.

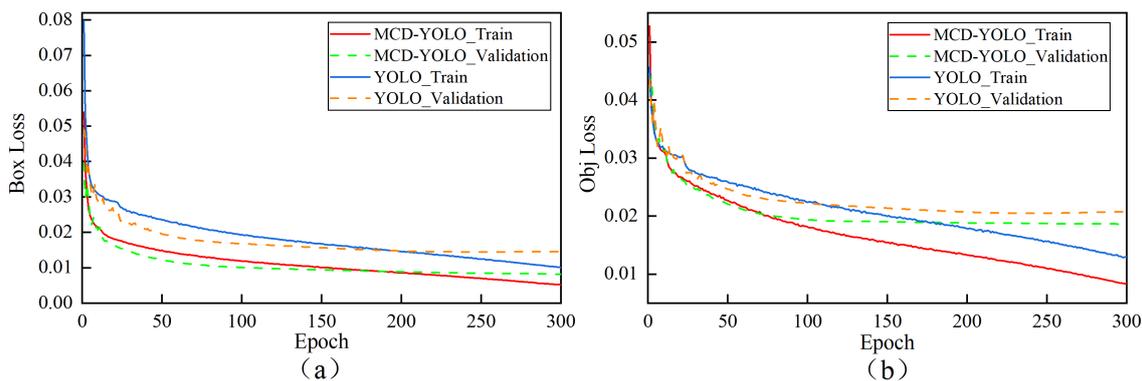


Figure 8. Comparing loss-change curves MCD-YOLOv5 and YOLOv5 training. The clump-detection task has only one category of clump, so the loss contains only localization loss and confidence loss: (a) variation in localization loss (Box Loss) with number of training rounds (epoch), (b) variation in confidence loss (Obj Loss) with number of training rounds (epoch).

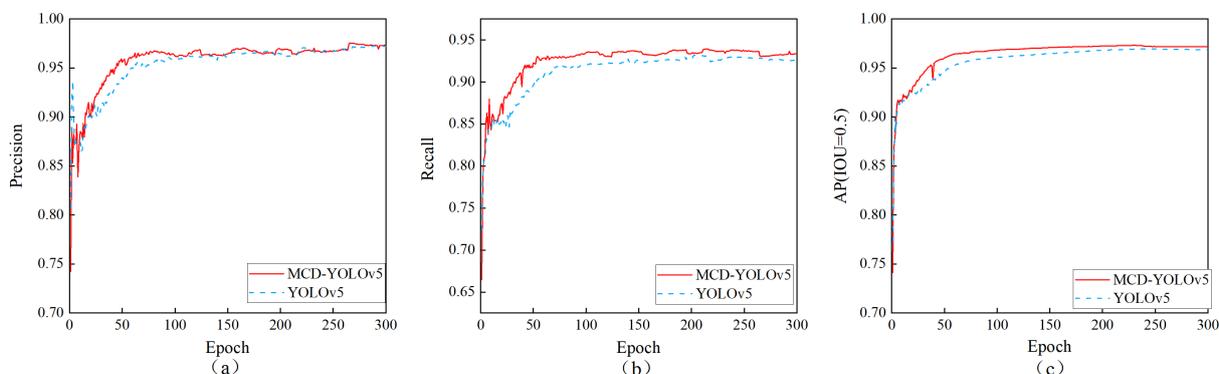


Figure 9. Variation curves of precision, recall, and AP as a function of the number of training rounds (epoch) for MCD-YOLOv5 and YOLOv5 on the validation set: (a) precision, and (b) recall, and (c) AP.

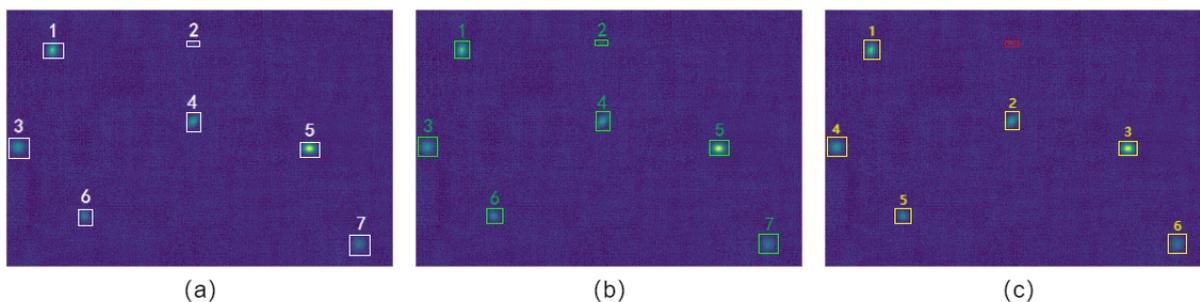


Figure 10. A typical example of detection results of MCD-YOLOv5 and YOLOv5. (a) Seven simulated clumps in the area. (b) MCD-YOLOv5 detected 7 clumps. (c) YOLOv5 detected 6 clumps. The red rectangular box in (c) represents the missed clump.

Table 4. Detection results of MCD-YOLOv5 and YOLOv5 on the test set.

Model	Precision	Recall	AP
MCD-YOLOv5	0.969	0.935	0.972
YOLOv5	0.969	0.910	0.956

4.2.3. Result of DPC

When MCD-YOLOv5 completes the detection, the center of mass of the detected clumps and their corresponding intensity information is inputted into DPC according to the

method mentioned in Section 3.2. DPC then classifies the centers of mass belonging to the same clumps into the same class. The results of DPC are shown in Figure 11. The differently colored dots represent the center of mass of a clump detected by MCD-YOLOv5, the red rhombus represents a clustering center, and the green square represents the location of the center of mass of a simulated clump. If the dots representing the detected clumps have the same color, these centers of mass are classified as one clump. Figure 11 shows that most of the cluster centers obtained through DPC can correspond with the centers of mass of the simulated clumps. This finding indicates that using intensity-based DPC can classify slices belonging to the same clumps into a single class in the direction of velocity, thereby ultimately realizing the detection of 3D clumps in PPV.

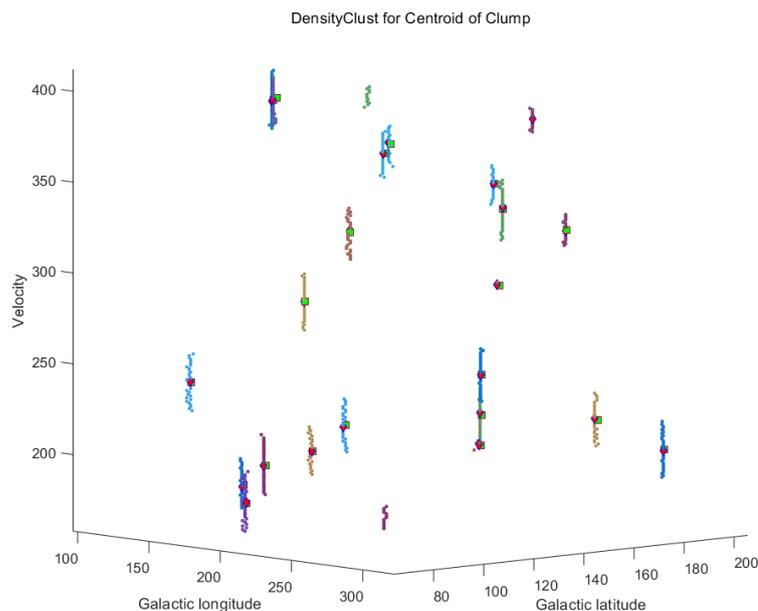


Figure 11. Display of the DPC results. The differently colored dots represent the center of mass of a clump detected by MCD-YOLOv5, the red rhombus represents a clustering center, and the green square represents the location of the center of mass of a simulated clump.

4.3. Detection Results of Second-Quadrant Synthesized Data

To test the detection performance of the proposed method based on MCD-YOLOv5 joint DPC, 100 second-Galactic-quadrant synthesized data (containing 10,000 simulated clumps) are selected for testing. To effectively evaluate the performance, we stipulate that if the centroid Euclidean distance between the detected clumps and the simulated clumps is not greater than two pixels (The minimum standard deviation of Gaussian distribution), the clumps are regarded as detected.

The 100 second-Galactic-quadrant synthesized data are intercepted along the velocity direction, yielding 42,400 slices of size 361×241 . The slices are inputted into the trained MCD-YOLOv5 to obtain the detection results. During MCD-YOLOv5 detection, the confidence is set to 0.6, i.e., the location information of the region is recorded only when the probability of being a clump exceeds 0.6. Then, the detection results of the same synthesized data are inputted into DPC to obtain the detection results of clumps in the synthesized data. Using MCD-YOLOv5 joint DPC, 9841 simulated clumps are correctly detected by matching with the simulated clump information tables, with a recall rate of 98.41%. Figure 12 shows the velocity-integrated intensity map of a detection result, where the white circles are the center of mass positions of the simulated clumps, and the red dots indicate the center of mass positions of the clumps that are detected and matched with the simulated clumps by MCD-YOLOv5 joint DPC. Figure 13 shows the integrated intensity map of the detected clump of second-quadrant synthesized data of different directions, the maximum

spectrum, and the average spectrum. As can be seen from the figure, the area detected by MCD-YOLOv5 joint DPC matches the characteristics of the simulated clumps.

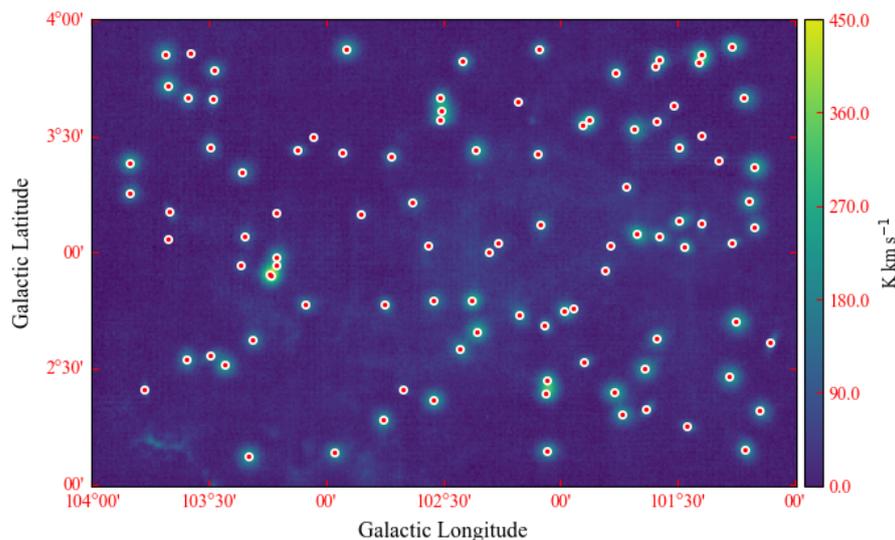


Figure 12. Velocity-integrated intensity maps of the detection result. The white circles are the center of mass positions of the simulated clumps, and the red dots indicate the center of mass positions of the clumps that are detected and matched with the simulated clumps by MCD-YOLOv5 joint DPC.

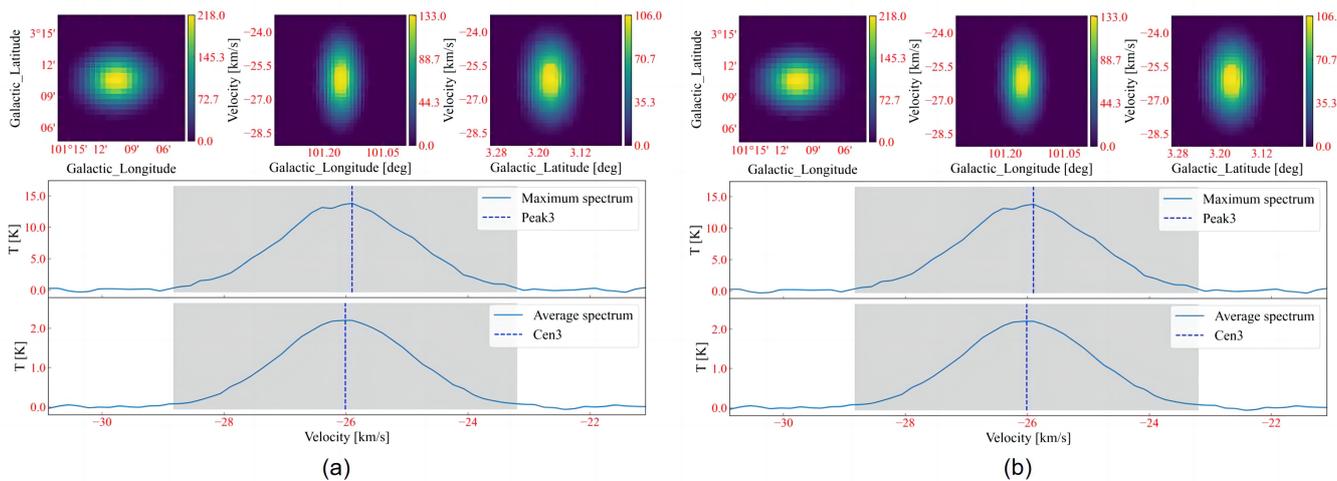


Figure 13. The information of the simulated clump is shown. (a,b) two examples of the integrated intensity map of the detected clumps of second quadrant synthesized data. The top three subplots offer the l–b, l–v, and b–v maps integration in three directions for a clump, while the middle and bottom of the figure show the peak spectrum and average spectrum of the clump, respectively.

To compare the performance of MCD-YOLOv5 joint DPC with the traditional clump-detection algorithms, we use FellWalker and ClumpFind on the same synthesized data. To improve the detection performance of FellWalker and ClumpFind, both algorithms are parameter-tuned according to the synthesized data to achieve better performance before detection. The detection parameters are determined by the value of the RMS of the MWISP and the experience of the users. When using MCD-YOLOv5 joint DPC, only two parameters need to be used, the threshold of density and distance in DPC. The detection parameters of FellWalker, ClumpFind, and MCD-YOLOv5 joint DPC are shown in Tables 5, 6, and 7, respectively. We determine whether each simulated clump is detected by matching the clump parameter tables with the output center-of-mass information of all three algorithms. Based on statistical results, 9841 clumps were detected by MCD-YOLOv5 joint DPC, 9770 by FellWalker, and 9631 by ClumpFind. When the second-Galactic-quadrant synthesized

data was detected using MCD-YOLOv5 joint DPC, 10,584 clumps were detected. Since the synthesized data includes the observational data background and simulated clumps, it needs to consider the number of clumps in the background data when calculating the error rate. The remaining 743 detected clumps were manually identified by manual verification, 495 clumps were falsely detected, and the false detection rate was 4.68%. Manual verification is based on the integrated intensity maps of the detected clump in different directions.

Table 5. FellWalker Parameters.

Parameters Name And Default Value
FELLWALKER.ALLOWEDGE = 1
FELLWALKER.CLEANITER = 1
FELLWALKER.FLATSLOPE = 2×RMS
FELLWALKER.FWHMBEAM = 2
FELLWALKER.MAXBAD = 0.05
FELLWALKER.MAXJUMP = 4
FELLWALKER.MINDIP = 1×RMS
FELLWALKER.MINHEIGHT = 3×RMS
FELLWALKER.MINPIX = 27
FELLWALKER.NOISE = 2×RMS
FELLWALKER.VELORES = 2

Table 6. ClumpFind parameters.

Parameters Name And Default Value
CLUMPFIND.ALLOWEDGE = 1
CLUMPFIND.DELTAT = 2×RMS
CLUMPFIND.FWHMBEAM = 2
CLUMPFIND.IDLAIG = 1
CLUMPFIND.MAXBAD = 0.05
CLUMPFIND.MINPIX = 27
CLUMPFIND.NAXIS = 3
CLUMPFIND.NOISE = 2×RMS
CLUMPFIND.TLOW = 3×RMS
CLUMPFIND.VELORES = 2

Table 7. MCD-YOLOv5 joint DPC parameters. minRho represents the minimum peak intensity of the clump, i.e., the point corresponding to this intensity can be used as the cluster center during the DPC clustering process. minRho can be set according to the intensity characteristics of the clumps in different regions. minDelta represents the minimum pixel distance to distinguish between two clumps. minDelta can be set according to the sparseness of the clump distribution.

Parameters Name	Explanation	Default Value
minRho	The minimum intensity of clump	[2, 5]×RMS
minDelta	The minimum pixel distance to distinguish between two clumps	4

In terms of the total number of detections, the method in this paper is slightly higher, but the number of parameters is greatly reduced and the setup is simpler. The recall rates based on MCD-YOLOv5 joint DPC, FellWalker, and ClumpFind detection results are shown in Table 8. Figure 14 demonstrates the variation of recall with flux and Peak Signal-to-Noise Ratio (PSNR). PSNR and flux of each simulated clump can be calculated by the clump parameter table. After the detection is completed, the detection results of the clumps are matched with the simulated clump parameter table. Clumps in the detection results that satisfy the matching rule are recorded as clumps detected correctly. The recall rate is statistically plotted according to the different intervals of PSNR and flux, which can show the detection performance of the algorithm for different PSNR and flux of clumps. It can be seen that MCD-YOLOv5 joint DPC is higher than the other two methods in most positions of the performance curve. In other words, the method in this paper requires only a small number of parameters to achieve the detection performance of the traditional algorithm with optimized parameters.

Table 8. Detection results of MCD-YOLOv5 joint DPC, FellWalker, and ClumpFind.

Method	Matched Clumps	Recall
MCD-YOLOv5 joint DPC	9841	98.41%
FellWalker	9770	97.70%
ClumpFind	9631	96.31%

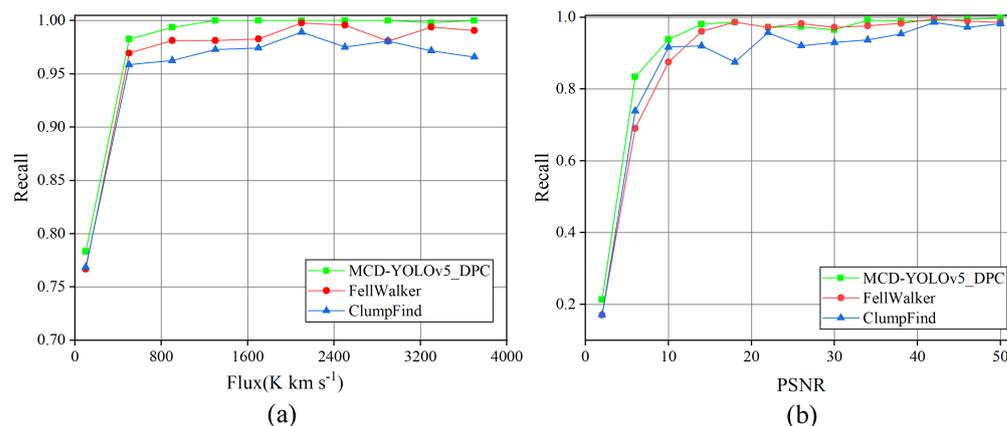


Figure 14. Variation of recall with flux and PSNR for ClumpFind, FellWalker, and MCD-YOLOv5 joint DPC: (a) Recall variation with flux, and (b) Recall variation with PSNR.

4.4. Detection Results of Observational Data

To verify the feasibility of MCD-YOLOv5 joint DPC on the observational data, we selected a region in the third Galactic quadrant of the $^{13}CO(1-0)$ line emission data obtained by MWISP for testing. The region selected ranges $180^\circ < l < 195^\circ$ and $-5^\circ < b < 5^\circ$, and the velocity ranges $-200 km s^{-1} < v_s < 200 km s^{-1}$. This region is divided into different data blocks at 1° grid spacing corresponding with the size of $121 \times 121 \times 2411$ pixels. To obtain annotation information, we used the LDC algorithm to detect this region and obtained the clumps and their corresponding parameter information. We selected the clumps among 200 data blocks with 487 clumps to obtain 8496 samples and their annotation information by the method mentioned in Section 4.2.1. The pixel size of the sample is 121×121 . Figure 15 shows the examples from the observational dataset and the annotation information on the samples. The dataset is divided in the same way as in Section 4.2.2. The training parameter is the default parameter of YOLOv5. To reduce the complexity of model training, we used Transfer Learning [35] when training MCD-YOLOv5 with the observational dataset. The trained model parameters on the synthesized dataset in Section 4.2.2 are used as the pre-training weights for the model training on the observational dataset. Figure 16 shows variation curves of loss-change, precision, and recall on the validation set

of the observational dataset. Figure 16a,b shows that MCD-YOLOv5 is not overfitted or underfitted when trained on the observational dataset. We selected some samples in the validation set and inputted them into the trained MCD-YOLOv5 for detection.

The detection results of MCD-YOLOv5 are shown in Figure 17. It can be seen that MCD-YOLOv5 can obtain the correct detection position of clump candidates in the 2D PP plane. After MCD-YOLOv5, we use DPC to obtain the final detection results of clumps in PPV space. The detection process and result outputs for observational data are identical to that for synthetic data, which is described in Section 4.3. Figure 18 demonstrates the velocity-integrated intensity maps of some of the detection results. It can be seen that the clumps are detected in a consistent location using both methods, MCD-YOLOv5 joint DPC and FellWalker. Figure 19 shows the integrated intensity map of the detected clump of observational data of different directions, the maximum spectrum, and the average spectrum. As can be seen from the figure, the area detected by MCD-YOLOv5 joint DPC matches the characteristics of the clumps.

By visual inspection, the detected candidate regions match the empirical criteria, demonstrating the viability of the proposed method in observational data. Moreover, the method in this paper does not detect spurious clumps caused by pure noise because there is no such type of target in the training labels. However, since the proposed method is based on supervised deep learning, the final performance is affected by the number of training samples and the accuracy of labeled information. During the experiments, there were also some missed detections of clumps due to incomplete training sets. When the background intensity is greater than the clump, the background will obscure the clumps, resulting in MCD-YOLOv5 missing the detection during the two-dimensional detection on the Galactic plane in the first stage. At the same time, a Gaussian-type clump occurs in multiple velocity channels, and the area of the clump is relatively small at the start and end of the velocity channel. MCD-YOLOv5 is prone to missed detection. Due to the loss of information at both ends, the resulting clump area will be smaller than the actual clump area after the DPC algorithm clusters. In this case, using the results of clustering to calculate the centroid will cause the result to be skewed from the actual position. Therefore, when compared with the simulated clump parameter table, the distance exceeds the size specified by the matching rule, so it is judged to be missed. In fact, the results of DPC clustering show that this type of clump is detected. In conclusion, the method in this paper is an attempt at deep learning for clump target detection in the PPV space.

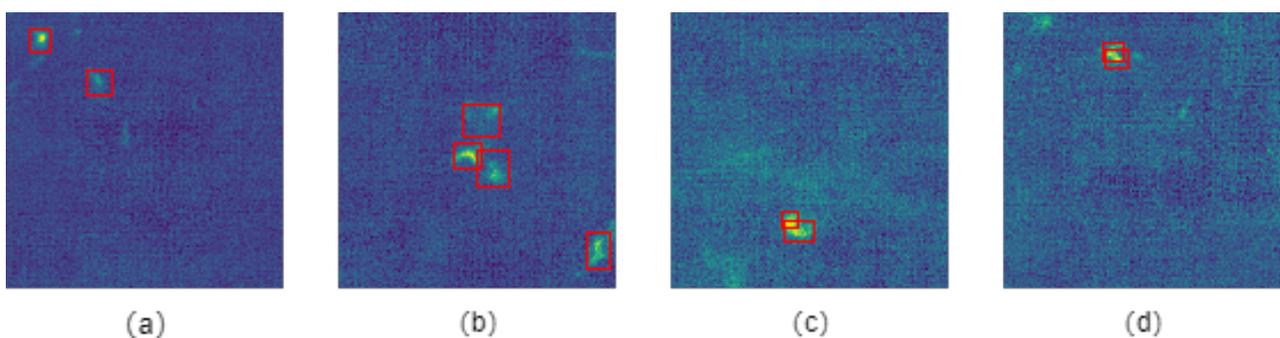


Figure 15. Examples from the MCD-YOLOv5 dataset and labeling the annotation information on the samples. (a–d) all show the labeling of the observational data slices in the third Galactic quadrant. The red-labeled box is annotation information.

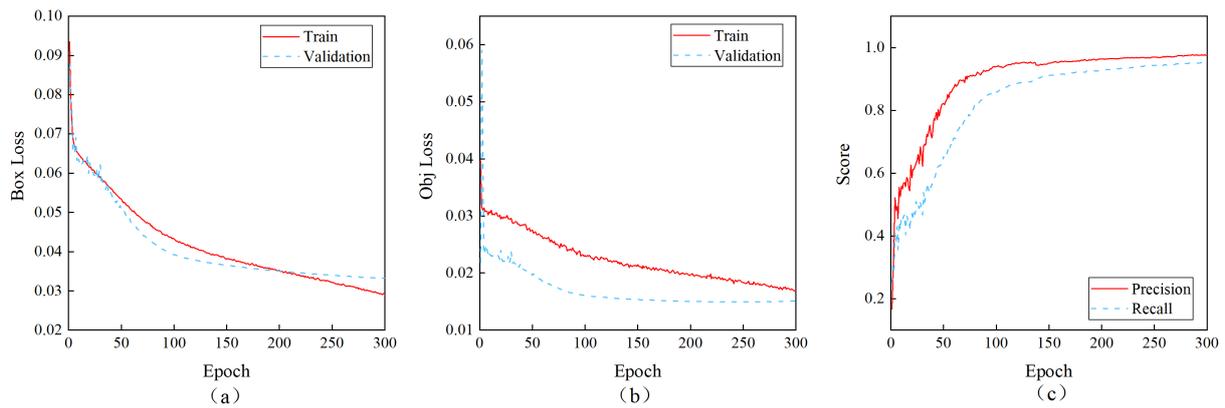


Figure 16. Variation curves of loss-change, precision, and recall on the validation set of the real dataset during MCD-YOLOv5 training: (a) variation in localization loss, (b) variation in confidence loss, (c) variation in precision and recall.

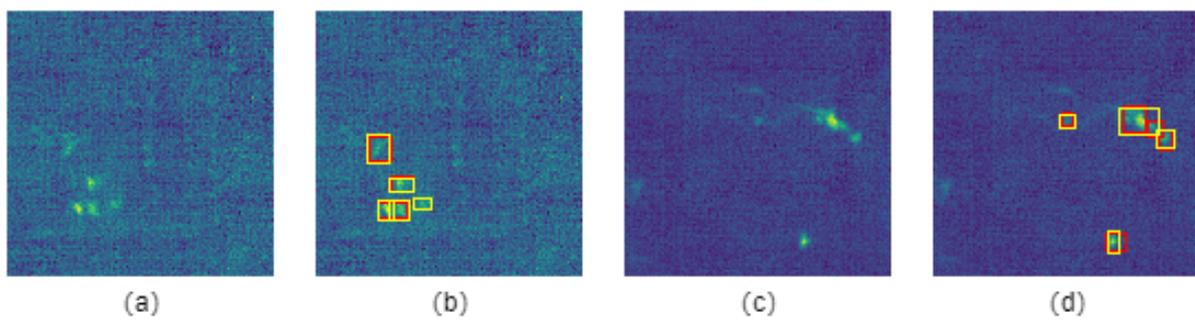


Figure 17. Detection results of MCD-YOLOv5 in observational data. (a,c) the slices generated by intercepting along the velocity channel of the observational data. (b,d) the detection results of MCD-YOLOv5. The red-labeled box is annotation information, the yellow-labeled box is the detection result of MCD-YOLOv5.

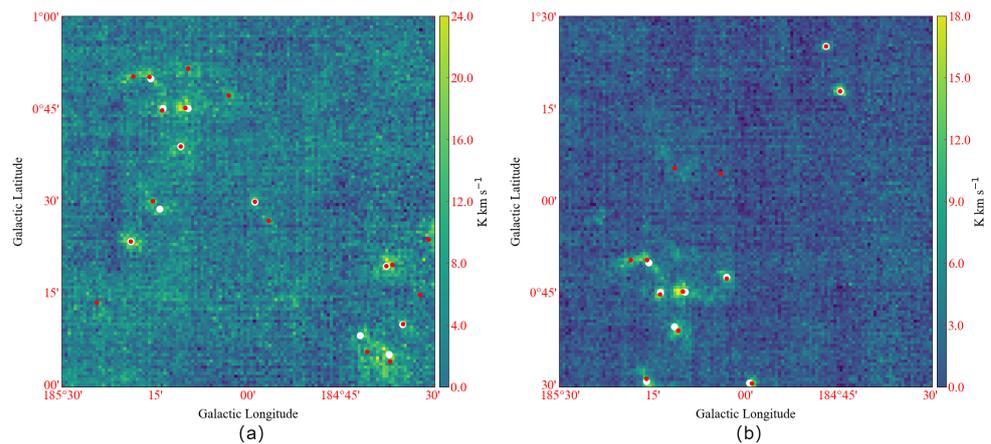


Figure 18. Velocity–integrated intensity maps of detection results. (a,b) velocity–integrated intensity maps of detection results of the two examples. The white circles are the center of mass positions of the clumps detected by MCD-YOLOv5 joint DPC, and the red dots indicate the center of mass positions of the clumps that are detected by FellWalker.

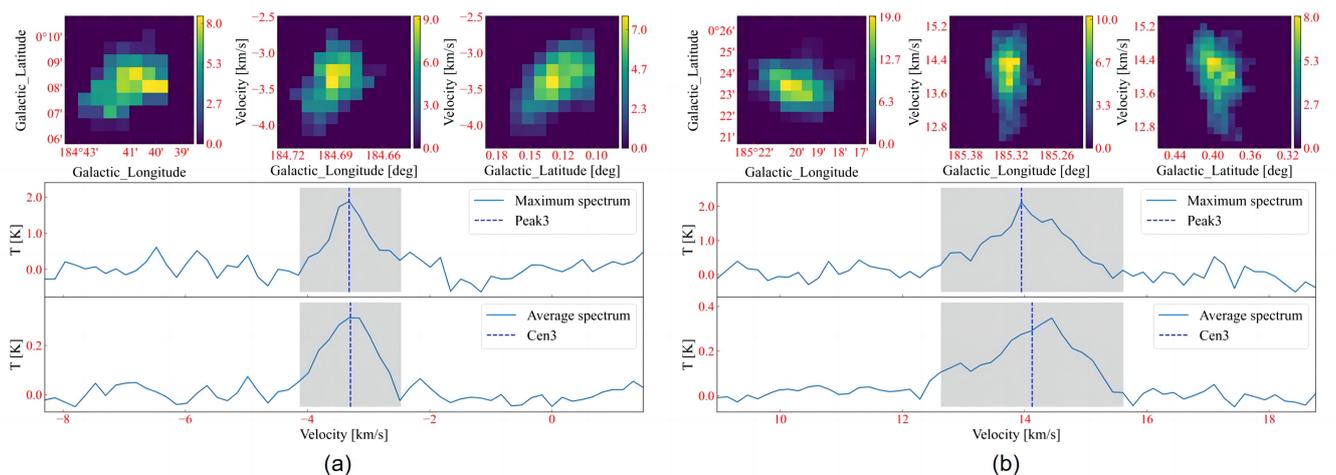


Figure 19. The information of the clump is shown. (a,b) two examples of the integrated intensity map of the detected clumps of observational data. The top three subplots offer the $l-b$, $l-v$, and $b-v$ maps integration in three directions for a clump, while the middle and bottom of the figure show the peak spectrum and average spectrum of the clump, respectively.

5. Summary

We propose a method based on an improved YOLOv5 joint DPC to realize the automatic detection of molecular clumps. Among them, YOLOv5 adds the CA module to the original backbone and modifies the loss function, which is used to improve the detection performance of small clumps. The intensity of the center of mass is considered in the process of DPC and is used to improve the clustering effect of the clumps. The algorithm localizes the molecular clumps on the Galactic plane by using MCD-YOLOv5 and combines the detection results in the velocity direction by using DPC to realize molecular-clump detection in PPV space. We constructed a large amount of synthetic data to test the detection performance of the algorithm. Experimental results show that the method proposed has fewer initial parameters, higher detection efficiency, and achieves the same detection performance as parameter-tuned FellWalker and ClumpFind. From the experiments with observational data, the method in this paper can also accurately find the positions where the clump candidates are located. Subsequently, the generalization ability of the model in observational data can be further improved by optimizing the quality of labels.

Author Contributions: Conceptualization, Z.-W.C. and Y.H.; methodology, S.Z. and Y.H.; software, J.-B.H. and X.-Y.L.; validation, C.L. and X.-Y.L.; formal analysis, J.-B.H. and C.L.; investigation, J.-B.H. and X.-Y.Z.; resources, Z.-W.C. and X.-Y.Z.; data curation, J.-B.H., C.L. and X.-Y.L.; writing—original draft preparation, J.-B.H. and Y.H.; writing—review and editing, X.-Y.Z. and S.Z.; visualization, J.-B.H., and X.-Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China U2031202.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The testing data that support the study are openly available in <https://github.com/SunetK/MCD-YOLOv5-joint-DPC> (accessed on 16 August 2023). Other datasets used or analyzed during the study are available from the corresponding author upon reasonable request.

Acknowledgments: We are grateful to the anonymous referee for their invaluable insights and comments, which enabled us to refine and enhance this work. This work is supported by the National Natural Science Foundation of China (grants Nos. U2031202). This research makes use of the data from the Milky Way Imaging Scroll Painting (MWISP) project, which is a multi-line survey in ^{12}CO , ^{13}CO , and C^{18}O along the northern Galactic plane with PMO-13.7m telescope. We are grateful to all the members of the MWISP working group, particularly the staff members at PMO-13.7m telescope, for their long-term support.

Conflicts of Interest: The authors declare no conflicts of interest.

Notes

¹ <https://github.com/SunetK/MCD-YOLOv5-joint-DPC>(accessed on 16 August 2023).

References

1. Heyer, M.; Dame, T.M. Molecular Clouds in the Milky Way. *Annu. Rev. Astron. Astrophys.* **2015**, *53*, 583–629. [[CrossRef](#)]
2. Williams, J.P.; Blitz, L.; McKee, C.F. The Structure and Evolution of Molecular Clouds: From Clumps to Cores to the IMF. *arXiv* **1999**, arXiv:astro-ph/9902246.
3. Stutzki, J.; Guesten, R. High Spatial Resolution Isotopic CO and CS Observations of M17 SW: The Clumpy Structure of the Molecular Cloud Core. *Astrophys. J.* **1990**, *356*, 513. [[CrossRef](#)]
4. Krumholz, M.R.; McKee, C.F.; Tumlinson, J. The Star Formation Law in Atomic and Molecular Gas. *Astrophys. J.* **2009**, *699*, 850–856. [[CrossRef](#)]
5. Zinnecker, H.; Yorke, H.W. Toward Understanding Massive Star Formation. *Annu. Rev. Astron. Astrophys.* **2007**, *45*, 481–563. [[CrossRef](#)]
6. Williams, J.P.; de Geus, E.J.; Blitz, L. Determining Structure in Molecular Clouds. *Astrophys. J.* **1994**, *428*, 693. [[CrossRef](#)]
7. Berry, D.S. FellWalker-A clump identification algorithm. *Astron. Comput.* **2015**, *10*, 22–31. [[CrossRef](#)]
8. Kirk, H.; Di Francesco, J.; Johnstone, D.; Duarte-Cabral, A.; Sadavoy, S.; Hatchell, J.; Mottram, J.C.; Buckle, J.; Berry, D.S.; Broekhoven-Fiene, H.; et al. The JCMT Gould Belt Survey: A First Look at Dense Cores in Orion B. *Astrophys. J.* **2016**, *817*, 167. [[CrossRef](#)]
9. Luo, X.; Zheng, S.; Huang, Y.; Zeng, S.; Zeng, X.; Jiang, Z.; Chen, Z. Molecular Clump Extraction Algorithm Based on Local Density Clustering. *Res. Astron. Astrophys.* **2022**, *22*, 015003. [[CrossRef](#)]
10. Chen, Q.; Wang, Y.; Yang, T.; Zhang, X.; Cheng, J.; Sun, J. You Only Look One-Level Feature. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, Virtual, 19–25 June 2021; pp. 13039–13048. [[CrossRef](#)]
11. Wang, J.; Song, L.; Li, Z.; Sun, H.; Sun, J.; Zheng, N. End-to-End Object Detection With Fully Convolutional Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, Virtual, 19–25 June 2021; pp. 15849–15858. [[CrossRef](#)]
12. Yan, B.; Peng, H.; Wu, K.; Wang, D.; Fu, J.; Lu, H. LightTrack: Finding Lightweight Neural Networks for Object Tracking via One-Shot Architecture Search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, Virtual, 19–25 June 2021; pp. 15180–15189. [[CrossRef](#)]
13. Kumar, A.; Rawat, Y.S. End-to-End Semi-Supervised Learning for Video Action Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, 18–24 June 2022; pp. 14680–14690. [[CrossRef](#)]
14. Liang, C.; Wang, W.; Zhou, T.; Yang, Y. Visual Abductive Reasoning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, 18–24 June 2022; pp. 15544–15554. [[CrossRef](#)]
15. Zhou, K.; Yang, J.; Loy, C.C.; Liu, Z. Conditional Prompt Learning for Vision-Language Models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, 18–24 June 2022; pp. 16795–16804. [[CrossRef](#)]
16. Kim, E.J.; Brunner, R.J. Star-galaxy classification using deep convolutional neural networks. *Mon. Not. R. Astron. Soc.* **2017**, *464*, 4463–4475. [[CrossRef](#)]
17. González, R.E.; Muñoz, R.P.; Hernández, C.A. Galaxy detection and identification using deep learning and data augmentation. *Astron. Comput.* **2018**, *25*, 103–109. [[CrossRef](#)]
18. Leung, H.W.; Bovy, J. Deep learning of multi-element abundances from high-resolution spectroscopic data. *Mon. Not. R. Astron. Soc.* **2019**, *483*, 3255–3277. [[CrossRef](#)]
19. Xie, J.; Bu, Y.; Liang, J.; Li, H.; Wang, X.; Pan, J. Improve the Search of Very Metal-poor Stars Using the Deep Learning Method. *Astron. J.* **2021**, *162*, 155. [[CrossRef](#)]
20. He, Z.; Qiu, B.; Luo, A.L.; Shi, J.; Kong, X.; Jiang, X. Deep learning applications based on SDSS photometric data: Detection and classification of sources. *Mon. Not. R. Astron. Soc.* **2021**, *508*, 2039–2052. [[CrossRef](#)]
21. Yi, Z.; Li, J.; Du, W.; Liu, M.; Liang, Z.; Xing, Y.; Pan, J.; Bu, Y.; Kong, X.; Wu, H. Automatic detection of low surface brightness galaxies from Sloan Digital Sky Survey images. *Mon. Not. R. Astron. Soc.* **2022**, *513*, 3972–3981. [[CrossRef](#)]
22. Cao, Z.; Yi, Z.; Pan, J.; Su, H.; Bu, Y.; Kong, X.; Luo, A. L-dwarf Detection from SDSS Images using Improved Faster R-CNN. *Astron. J.* **2023**, *165*, 184. [[CrossRef](#)]
23. Su, Y.; Yang, J.; Zhang, S.; Gong, Y.; Wang, H.; Zhou, X.; Wang, M.; Chen, Z.; Sun, Y.; Chen, X.; et al. The Milky Way Imaging Scroll Painting (MWISP): Project Details and Initial Results from the Galactic Longitudes of 25.°8–49.°7. *Astrophys. J. Suppl. Ser.* **2019**, *240*, 9. [[CrossRef](#)]
24. Rodriguez, A.; Laio, A. Clustering by fast search and find of density peaks. *Science* **2014**, *344*, 1492–1496. [[CrossRef](#)]
25. Matsubara, T. Analytic Minkowski functionals of the cosmic microwave background: Second-order non-Gaussianity with bispectrum and trispectrum. *Phys. Rev. D* **2010**, *81*, 083505. [[CrossRef](#)]
26. Li, Z.; Wang, Y.; Chen, K.; Yu, Z. Channel Pruned YOLOv5-based Deep Learning Approach for Rapid and Accurate Outdoor Obstacles Detection. *arXiv* **2022**, arXiv:2204.13699.

27. Darapaneni, N.; Kumar, S.; Krishnan, S.; Rajagopal, A.; Nagendra.; Paduri, A.R. Implementing a Real-Time, YOLOv5 based Social Distancing Measuring System for COVID-19. *arXiv* **2022**, arXiv:2204.03350.
28. Ewaidat, H.A.; Brag, Y.E. Identification of lung nodules CT scan using YOLOv5 based on convolution neural network. *arXiv* **2022**, arXiv:2301.02166.
29. Jain, S. Adversarial Attack on Yolov5 for Traffic and Road Sign Detection. *arXiv* **2023**, arXiv:2306.0607.
30. Hou, Q.; Zhou, D.; Feng, J. Coordinate Attention for Efficient Mobile Network Design. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021.
31. Yu, J.; Jiang, Y.; Wang, Z.; Cao, Z.; Huang, T. UnitBox: An Advanced Object Detection Network. *arXiv* **2016**, arXiv:1608.01471.
32. Wang, J.; Xu, C.; Yang, W.; Yu, L. A Normalized Gaussian Wasserstein Distance for Tiny Object Detection. *arXiv* **2022**, arXiv:2110.13389.
33. Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; Ren, D. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. *arXiv* **2019**, arXiv:1911.08287.
34. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man, Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
35. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.