Uncertainty-Guided Robotic Manipulation Through Variational Information Bottleneck in Imitation Learning

Qi Chen, Xinyang Ren, Weiyang Lin, *Member, IEEE*, and Chao Ye, *Member, IEEE*

***Abstract*—Traditional imitation learning methods typically rely on high-quality expert demonstrations and exhibit poor generalization when deployed in unfamiliar environments. A key limitation is their inability to effectively quantify and utilize epistemic uncertainty in the decision-making process. To address these limitations, this letter introduces a novel imitation learning framework that explicitly incorporates epistemic uncertainty estimation into policy learning. We leverage the Variational Information Bottleneck (VIB) to learn a compact and robust representation of the input data while simultaneously quantifying the uncertainty associated with each decision. Our method enables the model to generalize better to unseen scenarios and to make safer and more reliable decisions by reasoning about its own confidence in the predictions. Experimental results on various robotic manipulation tasks show that our method significantly improves performance compared to standard imitation learning methods, achieving better stability and adaptability.**

***Index Terms*—Imitation learning, planning under uncertainty, deep learning in grasping and manipulation.**

I. Introduction

ROBOTIC manipulation tasks are inherently challenging due to the complexity of environments, high degrees of uncertainty, and the need for robust and adaptive control strategies [1], [2]. Traditional robot control methods typically depend on pre-programmed rules and fixed task frameworks, assuming that the environment is known, static, and the tasks are deterministic. However, in the real-world, environments are often dynamic and uncertain, and tasks exhibit high levels of diversity and complexity. This mismatch between idealized assumptions and reality severely limits the adaptability and generalization capabilities of traditional methods. With recent advances in robotics and machine learning, data-driven methods have emerged as promising alternatives. Deep Reinforcement Learning (DRL) has shown great potential in enabling robots to learn complex control policies through interaction with uncertain and dynamic environments [3], [4], [5]. As a vital subset of DRL, imitation learning allows robots to acquire skills from expert demonstrations [6]. Compared to conventional DRL, imitation learning typically requires fewer interactions with the environment and enables faster policy learning.

However, directly applying imitation learning to robotic manipulation tasks still faces many limitations, and generalization capability and stability remain significantly limited [7], [8], particularly under environmental variations or noisy disturbances. When the input data is out of the training distribution, the learned policy model often generates unreliable action policies.

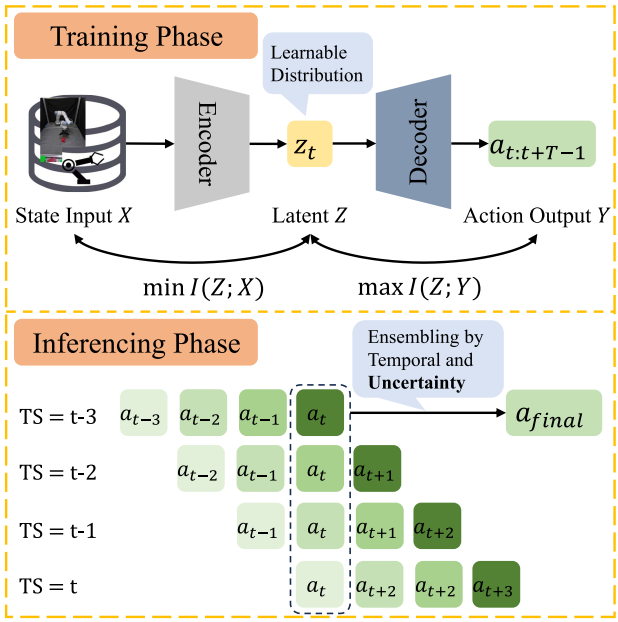


Fig. 1. **V**ariational **I**nformation **B**ottleneck Guided **I**mitation **L**earning **Control** (**VIB-ILCtrl**) framework. In the training phase, the VIB method is applied to maximize the mutual information between the latent variable Z and the action output Y ($I(Z;X)$) and minimize the mutual information between Z and the state input X ($I(Z;X)$). In the inferencing phase, the model generates multiple action predictions a_t at different time steps. These predictions are temporally ensembled based on epistemic uncertainty to produce the final action a_{final}.

Received 2 June 2025; accepted 16 September 2025. Date of publication 29 September 2025; date of current version 13 October 2025. This article was recommended for publication by Associate Editor M. Khoramshahi and Editor J. Kober upon evaluation of the reviewers' comments. This work was supported by the National Natural Science Foundation of China under Grant 62503134. *(Corresponding author: Chao Ye.)*

Qi Chen, Weiyang Lin, and Chao Ye are with the Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, Harbin 150001, China (e-mail: qich@stu.hit.edu.cn; wylin@hit.edu.cn; yechao@hit.edu.cn).

Xinyang Ren is with the School of Electrical and Information Engineering and the Yunnan Key Laboratory of Unmanned Autonomous Systems, Yunnan Minzu University, Kunming 650504, China (e-mail: rxy0424@gmail.com).

The code is available at https://github.com/Dear-Chen/VIB-ILCtrl.

This article has supplementary downloadable material available at https://doi.org/10.1109/LRA.2025.3615539, provided by the authors.

Digital Object Identifier 10.1109/LRA.2025.3615539

Furthermore, the phenomenon of compounded errors, where small initial mistakes during execution are amplified over time, can lead to significant deviations from the desired behavior and ultimately result in task failure [9], [10]. Traditional imitation learning methods usually lack uncertainty assessment mechanisms, limiting the model's ability to identify reliable prediction boundaries. As a result, their models fail to detect and recover from incorrect actions in time, resulting in poor real-world performance. This highlights the importance of integrating uncertainty estimation into imitation learning, enabling models to identify unreliable predictions and improve robustness in complex environments.

In this letter, we propose **V**ariational **I**nformation **B**ottleneck Guided **I**mitation **L**earning **Con**t**r**o**l** (**VIB-ILCtrl**), a novel imitation learning control framework based on the Variational Information Bottleneck (VIB). Our method learns compact and informative representations from expert demonstrations while explicitly modeling epistemic uncertainty in policy predictions. A key innovation is the use of a dynamic weighting mechanism that assigns lower confidence to uncertain actions, preventing the model from over-trusting unreliable decisions. This strategy improves generalization, robustness, and interpretability by prioritizing confident predictions and suppressing uncertain ones. We provide both the theoretical foundation and practical implementation of VIB-ILCtrl, followed by comprehensive experiments that validate its effectiveness in handling error propagation, distributional shifts, and uncertain environments.

The main contributions of this letter are as follows:

1) We are the first to incorporate the Variational Information Bottleneck method into the imitation learning framework for epistemic uncertainty estimation.
2) We propose a novel uncertainty-guided mechanism to mitigate compounded errors, effectively reducing the risk of incorrect decisions over subsequent steps.
3) We present an uncertainty-aware framework that integrates adaptive imitation learning with risk assessment, enabling robots to make efficient, accurate, and safe decisions in dynamically disturbed environments.

The remainder of this letter is organized as follows: Section II reviews related work, Section III details the proposed VIB-ILCtrl method, Section IV presents experimental settings and results, Section V concludes the paper and outlines future research directions.

II. Related Work

A. Imitation Learning

Imitation learning has gained significant attention in robotics for its capacity to learn complex tasks by observing and mimicking human or expert demonstrations. Early methods of imitation learning primarily focused on Behavior Cloning (BC), where a policy is learned by directly mapping observations to actions based on expert demonstrations [11]. However, BC suffers from compounded errors over time and limited generalization to new situations [12].

To address these limitations, researchers have investigated various advanced approaches. One prominent method is Inverse Reinforcement Learning (IRL) [13], which infers the reward function from expert demonstrations and then optimizes the policy to maximize this inferred reward. Additionally Adversarial Imitation Learning (AIL) [14], [15], use adversarial training to make an agent's behavior indistinguishable from that of an expert by matching their action distributions. However, these methods often require large-scale high-quality demonstrations and struggle with generalization to unseen scenarios. Despite these advancements, challenges remain, particularly in the context of handling dynamic environments, limited demonstration data, and safety-critical tasks.

B. Uncertainty Estimation

Estimating model uncertainty is essential for enhancing the reliability and robustness of predictions in machine learning. Diverse methods have been established to quantify uncertainty, each with its own strengths and limitations.

Bayesian methods [16], [17] assess uncertainty by formulating posterior distributions over model parameters. The primary advantage of Bayesian methods is their theoretical reliability, which provides direct measures of parameter uncertainty. However, Bayesian inference frequently involves high computational expenses, particularly in deep learning, where efficient approximate inference remains a significant challenge.

Ensemble methods [18], [19], [20] enhance performance and estimate uncertainty by combining predictions from multiple models. Deep Ensembles [18], as a basic yet effective strategy, reduce variance and bias by averaging predictions from models trained with different initializations or subsets of data. An et al. [19] introduce an uncertainty-based offline RL method that leverages an ensemble of Q-networks (SAC-N) to enhance the penalty on out-of-distribution state-action pairs, thereby significantly improving algorithmic performance. Nevertheless, ensemble methods require additional computational resources to train and store multiple models, leading to significantly increased computational complexity, training time, and memory consumption.

Random Network Distillation (RND) [21] estimates state uncertainty by comparing the outputs of a fixed target network and a trainable predictor network. Despite its effectiveness in capturing novelty in environments, RND is highly sensitive to the initialization of the two neural networks. Since the target network's weights remain fixed while the predictor network is trained from scratch, the results may vary significantly depending on the initial parameter settings. This sensitivity can lead to inconsistent learning trajectories and performance, raising challenges for its stability and general applicability.

C. Variational Information Bottleneck

The Information Bottleneck (IB) [22] theory is an information-theoretic framework aimed at identifying a data representation that retains the most relevant information for a specific task while simultaneously discarding extraneous details from the original dataset. Derived from the IB theory, the Variational Information Bottleneck (VIB) [23], [24] framework attempts to achieve an optimal balance between compressing

input data and retaining the essential relevant information for a given task.

Studies on VIB contain various applications, including imitation learning, inverse reinforcement learning, domain generalization, representation learning, and manipulation tasks. Hsu et al. [25] propose the utilization of VIB to regularize the information streams in vision-based manipulators, with a particular emphasis on tasks centered around hand movements. Peng et al. [14] introduce the Variational Discriminator Bottleneck (VDB), which applies constraints to information flow within discriminators, resulting in enhanced performance of adversarial learning algorithms. Du et al. [26] tackle the challenges of prediction uncertainty and domain shift in domain generalization models through the utilization of VIB. Furthermore, Kudo et al. [27] propose the Flexible Variational Information Bottleneck (FVIB), which attains optimal compression-prediction trade-offs across a continuous range of Lagrange multipliers via a single training run and markedly improves model calibration.

III. Method

A. Preliminaries

Formally, agents in imitation learning are modeled as a Markov Decision Process (MDP) defined by a tuple $\langle S, A, T, R\rangle$, where the parameters represent the state space, action space, transition function, and reward function. Imitation Learning relies on a dataset $\mathcal{D}$ of expert demonstrations to learn a policy that mimics the expert's behavior. The dataset $\mathcal{D}$ includes sequential trajectories τ, each of which is typically composed of state-action pairs collected from the expert's interactions with the environment. Specifically, each trajectory τ is represented as:

$$\tau = (s_1, a_1), (s_2, a_2), \ldots, (s_T, a_T) \tag{1}$$

where T is the length of the trajectory. Here, $a_t \in A$ and $s_t \in S$ denote the action and state at current time step t. In this work, the state s_t includes both image observation I_t and robot proprioception p_t, such that: $s_t = (I_t, p_t)$. The proprioception data include joint angles, end-effector positions, and other relevant sensor readings.

The goal of imitation learning is to minimize the discrepancy between the learned policy function $\pi_\theta(a_t|s_t)$ parameterized by θ and the expert policy $\pi^*(a_t|s_t)$. Typically, it's formulated as an optimization problem:

$$\mathcal{L}_{\text{IL}} = - \, mathbbE_{(s_t,a_t)\sim\mathcal{D}}[\log \pi_\theta(a_t|s_t)]. \tag{2}$$

B. Variational Information Bottleneck

The Variational Information Bottleneck (VIB) [23], [24] framework is a powerful method that incorporates the principles of the Information Bottleneck (IB) [22] method into a variational setting and aims to improve the robustness and generalization of machine learning models. It treats the input data X and the target variable Y as random variables and seeks to learn a stochastic mapping from X to a latent representation Z that captures the essential information for predicting Y. The key idea is to maximize the mutual information $I(Z;Y)$ between Z and Y while constraining the mutual information $I(Z;X)$ between Z and X. This constraint serves as a bottleneck, forcing the model to retain only the most relevant information. Mathematically, the network parameterized by θ and the VIB objective can be formulated as:

$$\max I(Z;Y|\theta) - \beta I(Z;X|\theta) \tag{3}$$

where β is a Lagrange multiplier that controls the trade-off between retaining information about the input and the target.

C. Epistemic Uncertainty Estimation

Quantifying epistemic uncertainty is crucial for modern machine learning performance, especially in dynamic environments where models must recognize and calibrate uncertainty on out-of-distribution data to avoid overconfident errors. The VIB [23], [24] framework provides a novel perspective on epistemic uncertainty quantification. In this framework, the rate R, defined as the Kullback-Leibler (KL) divergence between encoder $e_\theta(z|x)$ and the learnable marginal distribution $m_\theta(z)$ in the latent space, serves as an effective uncertainty indicator:

$$R = \mathrm{KL}[\mathrm{e}_\theta(\mathrm{z}|\mathrm{x}) \parallel \mathrm{m}_\theta(\mathrm{z})] \tag{4}$$

Here, $e_\theta(z|x)$ encoders input x into the latent representation z, and $m_\theta(z)$ is a learnable Gaussian prior, $m_\theta(z) \sim \mathcal{N}(\mu, \delta^2)$, with trainable mean μ and variance δ^2.

This measure reflects the difference between the latent representation of the input sample x and the standard representations learned from the training data. A low rate R indicates that the input aligns well with the training distribution, whereas a significant increase in R suggests a deviation from it, enabling the model to effectively assess the epistemic uncertainty.

D. The Training of VIB-ILCtrl

Building on VIB, we propose VIB-ILCtrl, a novel framework that integrates VIB into imitation learning. We design the imitation policy $\pi_\theta(a_{t:t+T-1}|s_t)$ of VIB-ILCtrl following ACT [12], which learns action chunks over a horizon T rather than individual actions. This method allows the model to generate more dependable sequences of actions via a temporal ensemble. To integrate VIB into the imitation learning framework, we first refine the standard imitation learning policy into three separate components:

- Encoder Network: $p_\theta(z_t|s_t)$, which maps the state s_t to the latent representation z_t.
- Decoder Network: $q_\theta(a_{t:t+T-1}|z_t)$, which generates actions $a_{t:t+T-1}$ based on the latent representation z_t.
- Learnable Marginal Distribution: $m_\theta(z_t)$, which represents the prior distribution over the latent representation z_t.

According to the VIB framework, we define the following composite loss function:

$$\begin{aligned}
\mathcal{L}_{\text{reconst}} &= - \, mathbbE_{(s_t,a_{t:t+T-1})\sim\mathcal{D}}[\log q_\theta(a_{t:t+T-1}|z_t)] \\
\mathcal{L}_{\text{rate}} &= \mathrm{KL}[p_\theta(z_t|x_t) \parallel m_\theta(z_t)] \\
\mathcal{L}_{\text{total}} &= \mathcal{L}_{\text{reconst}} + \beta\mathcal{L}_{\text{rate}}
\end{aligned} \tag{5}$$

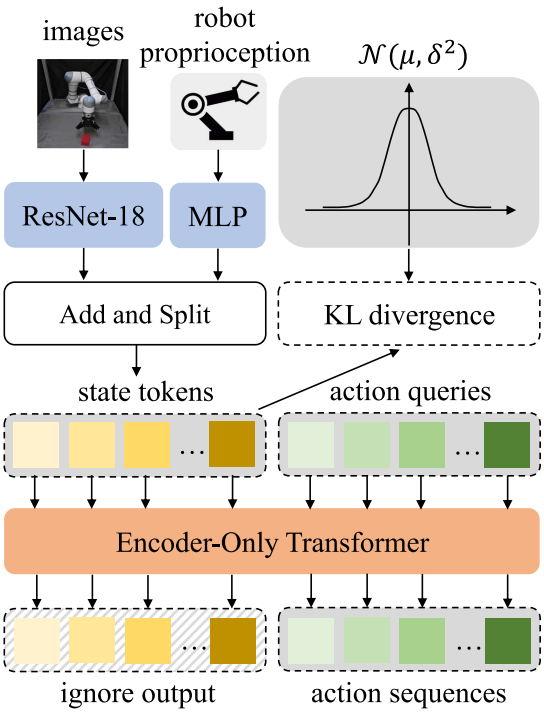


Fig. 2. The Variational Information Bottleneck Guided Imitation Learning Control (VIB-ILCtrl) method architecture.

The total loss function in Eq. (5) consists of two key components: a reconstruction loss $\mathcal{L}_{\text{reconst}}$ and a rate loss $\mathcal{L}_{\text{rate}}$. The hyperparameter β, which corresponds to the Lagrange multiplier in Eq. (3), serves to balance these two terms. The reconstruction loss ensures accurate recovery of ground-truth actions from the latent representation, while the rate loss aligns the latent distribution z_t with the learnable marginal distribution $m_\theta(z_t)$.

The network architecture is detailed in Fig. 2. The encoder network $p_\theta(z_t|s_t)$ receives as input environmental images I_t and robot state information p_t, processing them through two parallel pathways: a ResNet-18 [28] for image feature extraction and a Multi-Layer Perceptron (MLP) for robot proprioceptive data encoding. Following the extraction of visual and proprioceptive features, the system fuses them by addition and subsequently splits the fused features along the last dimension into multiple state tokens, each encoding integrated state information. The resulting multiple state tokens are then passed to the subsequent processing stage.

A key component of our method is computing the Kullback-Leibler (KL) divergence between the multiple state tokens and the latent representation's marginal distribution, which is defined as a learnable Gaussian distribution. This step aims to enhance the system's generalization capabilities and assess epistemic uncertainty. By minimizing the $\text{KL}[p_\theta(z_t|x_t) \parallel m_\theta(z)]$, the model effectively balances the compression of unnecessary information while keeping key features. Meanwhile, the result, which is also defined as rate R, is used to measure the epistemic uncertainty.

In VIB-ILCtrl, the multiple state tokens are fed into the decoder $q_\theta(a_{t:t+T-1}|z_t)$, an Encoder-Only Transformer architecture equipped with learnable action queries [29], to generate output action sequences. Unlike traditional transformer models, which consist of both encoder and decoder components, our architecture employs a single encoder to directly map the state tokens to the action commands. Our method generates action sequences using parallel action queries rather than autoregressive models. Given that our generation process does not require consideration of causal relationships between states and actions, we choose an Encoder-Only architecture to achieve this.

E. The Inference of VIB-ILCtrl

Our method estimates current state uncertainty using Eq. (4). To ensure safe and reliable operation during practical inference, the model dynamically adjusts the action generation based on this epistemic uncertainty estimate. Specifically, a threshold γ is used to decide the final action. When the estimated epistemic uncertainty exceeds the threshold γ, the robot terminates the current task and executes a predefined safety action. Examples of such actions include returning to the starting position or halting operations entirely. Conversely, when the epistemic uncertainty falls below the threshold, the system continues with normal action generation to ensure efficient and reliable task execution. In practice, we recommend setting the threshold γ around the mean plus twice the standard deviation of the training data's uncertainty, with minor adjustments based on the specific application.

In the inference process of the ACT [12] method, action chunks are learned and subsequently smoothed using an exponential moving average (EMA) to generate the final output. This smoothing mechanism facilitates continuous and stable transitions between sequential actions. However, we observe that this method can be sub-optimal in certain scenarios, particularly when unreliable recent actions are assigned higher weights compared to reliable past actions.

To address this issue, we incorporate epistemic uncertainty estimation from VIB by assigning an uncertainty measure u_t to each action a_t, and adaptively adjusting its weight based on this value. The final action output is computed as a weighted average based on the temporal ensemble and uncertainty of each action:

$$a_{final} = \sum_{t=0}^{T-1} w_{t-i} a_{t-i}$$

$$w_t = \frac{\exp(u_t * m * i)}{\sum_{i=0}^{T-1} \exp(u_i * m * i)} \tag{6}$$

where w_t represents the weight. Here m is used to control the trade-off between temporal ensemble and uncertainty weighting, ensuring that more reliable past actions are not overshadowed by potentially unreliable recent actions. To the best of our knowledge, we are the first to utilize epistemic uncertainty measures to evaluate the reliability of actions and guide action generation. In actual training, we find that the uncertainty rate varies across different datasets and the Lagrange multiplier. So in the inference phase, we divide the uncertainty rate by the mean uncertainty rate of the data set.

F. Design Rationale and Comparison With Alternatives

Although diffusion models [30] offer powerful generative capabilities, the inference latency exceeds the real-time

requirement of robotic systems. In contrast, VIB enables efficient, single-pass uncertainty estimation with minimal computational overhead, which is a critical advantage for real-time deployment. While VIB shares a close theoretical relationship with Variational Autoencoders (VAEs) [31], differs in objective. VAEs aim to reconstruct inputs, whereas VIB seeks to preserve task-relevant information under a compression constraint. Furthermore, VIB provides an information-theoretic, end-to-end trainable framework that balances predictive accuracy and representation robustness, improving generalization under environmental perturbations.

IV. Experiments

In this section, we present a comprehensive evaluation of the proposed VIB-ILCtrl method.

A. Experimental Setup

1) Experimental Environment: In the simulated experiments, we evaluate the proposed method in an ALOHA environment on two bimanual manipulation tasks: a transfer task and an insertion task. In the transfer task, the robot uses its right arm to grasp a red cube and then transfers it to the left arm. In the insertion task, the right arm grasps a red peg and inserts it into a blue socket held by the left arm. For each task, we use two types of demonstrations: one generated by scripted policies and the other collected from human, with 50 trajectories per type for training.

For the real-world experiment, we deploy the method on a 6-degree-of-freedom industrial collaborative robotic arm equipped with an electric gripper. We design a grasping task, in which the robot is required to accurately grasp a red block placed on a table. A total of 50 expert trajectories are collected for this task.

During evaluation, each task is tested under both normal conditions and perturbed settings involving Gaussian noise or color jitter. These perturbations are designed to simulate realistic, out-of-distribution conditions that closely resemble practical deployment scenarios. Gaussian noise with zero mean and standard deviation $\delta \in \{0, 0.05, 0.1\}$ is added to the input images, with each noise level applied with equal probability $1/3$. Color jitter perturbations (brightness, contrast, saturation, and hue) are applied with a strength of 0.2 per channel, also with a 50% probability. In simulation, each method is evaluated over 100 trials to ensure statistical reliability. In real-world tasks, considering operational complexity and cost constraints, we conduct 50 trials per method and compute the success rate accordingly.

2) Baseline: To evaluate the effectiveness of our proposed method, we compare it against three baseline methods:

- BC-ConvMLP [32] : A classical and widely adopted baseline in the imitation learning paradigm. It uses a convolutional neural network to process visual inputs and a multi-layer perceptron to map states to actions.
- Diffusion Policy [30] : An innovative algorithm for generating robot action, which represents a robot's policy as a conditional denoising diffusion process.
- ACT [12] : A state-of-the-art imitation learning framework that leverages a transformer-based architecture to learn action chunks, followed by an exponential moving average (EMA) mechanism for smoothing action predictions.

B. Results Analysis

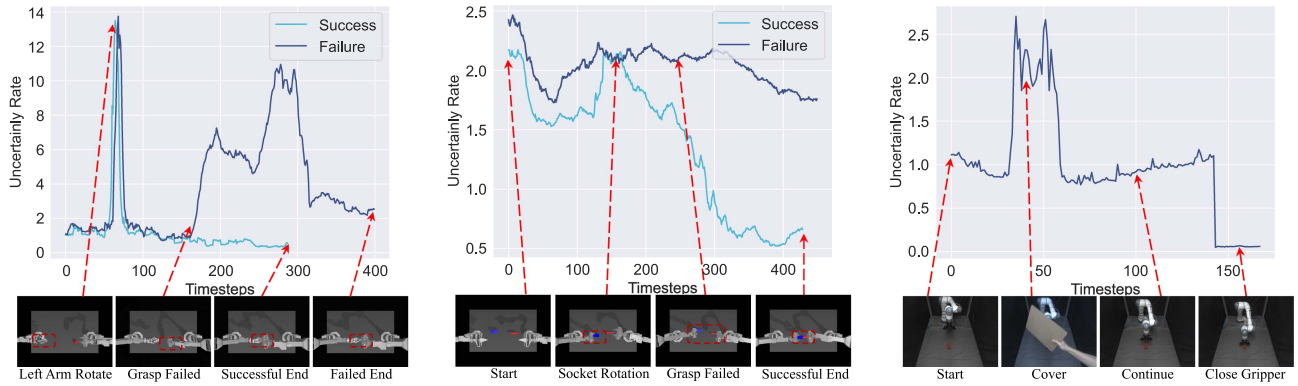
1) Simulated Experiments: Table I summarizes the simulated experimental results by comparing the success rates of different approaches in performing transfer and insertion tasks under normal conditions, Gaussian noise, and color jitter. BC-ConvMLP achieves consistently low success rates across all test settings, ranging from 0% to 4%, indicating limited generalization and robustness. The Diffusion Policy method performs better than BC-ConvMLP on the Transfer task. Specifically, it achieves a success rate of 79% on scripted data, but this drops significantly to 36% when trained on human demonstrations. Under Gaussian noise or color-jittered conditions, the performance deteriorates further. In contrast, Diffusion Policy shows poor performance on the insertion task across both data sources. ACT achieves higher performance than both BC-ConvMLP and Diffusion Policy. It reaches a success rate of 89% for the scripted transfer task under normal conditions. However, its performance degrades substantially across other settings, achieving only 46%, 12%, and 26% on the transfer (human), insert (scripted), and insert (human) tasks. Its performance also declines notably under perturbed conditions. Our proposed method, VIB-ILCtrl, achieves the best performance in terms of accuracy and robustness among all evaluated methods. Under normal conditions, it attains success rates of 96%, 81%, 42%, and 43% on the four tasks. Even under Gaussian noise or color jitter, although the performance decreases slightly, VIB-ILCtrl still significantly outperforms all baselines.

Fig. 3(a) shows the epistemic uncertainty rate during a successful and a failed rollout of a transfer task. Initially, the uncertainty remains low in both cases. Around the 70th timestep, a notable spike occurs in both cases due to the rotation of the left robotic arm. In the failed case, uncertainty rises sharply at the 170th step when the gripper fails to grasp the red block and remains high. In contrast, the successful trajectory completes the task at the 280th step with only moderate uncertainty. By the 400th timestep, although the robot pose in the failed case is similar to that of the successful one, the absence of the red block results in persistently elevated uncertainty. Fig. 3(b) further illustrates the uncertainty rate in a successful and a failed Insertion Task. In the successful case, uncertainty increases when the left socket is rotated, but drops after the successful grasp. In contrast, the failed task maintains a consistently high uncertainty level throughout. These observations indicate that the proposed uncertainty estimation effectively detects out-of-distribution or erroneous states.

2) Real-World Experiments: Table III demonstrates the real-world experimental results. Our method achieves an overall success rate of 80%, significantly outperforming the ACT baseline, which attains an average success rate of 54%. Notably, even under perturbed conditions involving Gaussian noise and color jitter, our method maintains a relatively high performance. In

TABLE I
SUCCESS RATE FOR THE SIMULATED TASKS, COMPARING OUR METHOD TO THE BASELINES

	Transfer (Scripted)			Transfer (Human)			Insertion (Scripted)			Insertion (Human)		
	Normal	Gaussian	Color	Normal	Gaussian	Color	Normal	Gaussian	Color	Normal	Gaussian	Color
BC-ConvMLP	4	0	0	1	0	0	1	0	0	0	0	0
Diffusion Policy	79	72	75	36	30	33	10	5	6	6	8	6
ACT	88	86	85	46	34	40	12	6	10	26	14	24
Ours	**96**	**95**	**95**	**81**	**79**	**76**	**42**	**40**	**39**	**43**	**37**	**40**
ACT w/Uncertainty	86	89	86	48	40	43	17	10	16	31	20	28
ACT w/RND	90	86	88	50	47	42	20	15	14	28	17	30
Ours wo/Uncertainty	94	93	94	71	72	71	33	23	29	33	33	31
Ours r/RND	**96**	94	**95**	78	77	**76**	38	35	37	40	**37**	38



(a) A successful episode and a failed episode of Transfer task in simulation

(b) A successful episode and a failed episode of Insertion task in simulation

(c) A successful episode with occlusion of Grasp task in real-world

Fig. 3. Some examples of the uncertainty rate during the task execution. Plot shows uncertainty rate over time, with corresponding key frames labeled with RGB images, red arrows, and red labels.

TABLE II
INFLUENCE OF THE LAGRANGE MULTIPLIER β

	Transfer (Scripted)			Transfer (Human)			Insertion (Scripted)			Insertion (Human)		
	Normal	Gaussian	Color	Normal	Gaussian	Color	Normal	Gaussian	Color	Normal	Gaussian	Color
$\beta = 0.1$	30	21	30	16	11	10	2	1	2	0	0	0
$\beta = 0.01$	33	29	31	54	44	42	14	5	9	3	3	3
$\beta = 0.001$	**96**	94	94	50	50	48	25	24	23	27	20	23
$\beta = 0.0001$	**96**	**95**	**95**	**81**	**79**	**76**	**42**	**40**	**39**	**43**	**37**	**40**

contrast, the performance of all baseline methods drop considerably under the same conditions, with Diffusion Policy and BC-ConvMLP exhibiting even worse performance and robustness. These results indicate that incorporating uncertainty estimation into the action generation process effectively enhances system robustness and reliability. Notably, BC-ConvMLP exhibits significant performance degradation in long-horizon tasks, consistent with the known limitation of single-step behavioral cloning: error accumulation over time renders it unsuitable for precise, sequential control [12].

To validate our method's ability to detect uncertainty, we conducted a real-world grasping environment under occlusion, as shown in Fig. 3(c). Between timesteps 0 and 25, the uncertainty rate remains low, showing that the robot can accurately recognize the target object and perform corresponding operations. However, with the introduction of an occlusion at approximately the 30th timestep, the uncertainty rate increases significantly. To prevent the robot from performing unreliable actions under such high uncertainty, the robot decides to pause its operation based on a preset threshold γ for safety. Subsequently, after the occlusion is removed (about 60th timestep), uncertainty gradually declines and stabilizes, allowing the robot to resume grasping. At about the 140th timestep, the gripper begins to close.No further actions are required, and the uncertainty rate

TABLE III
SUCCESS RATE FOR THE REAL-WORLD TASKS, COMPARING OUR METHOD TO THE BASELINES

	Grasp			
	Normal	Gaussian	Color	Avegare
BC-ConvMLP	2	4	2	2.7
Diffusion Policy	38	26	20	28.0
ACT	66	50	46	54.0
Ours	**84**	**76**	**80**	**80.0**

TABLE IV
MORE GENERALIZATION EVALUATION RESULTS

	Environmental Variations			
	Normal	Camera Shake	Unseen Block	Occlusion
Ours	84	76	62	82

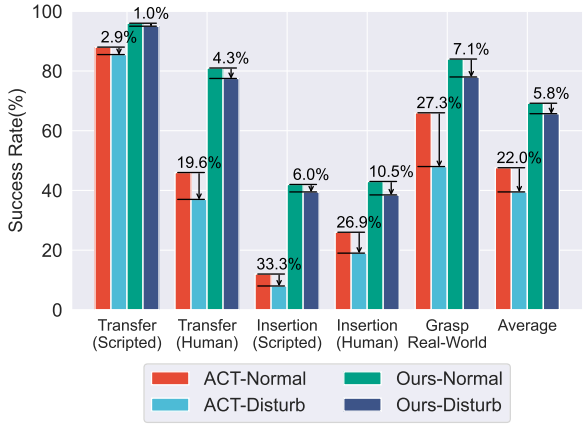


Fig. 4. Compare our method to ACT based on the success rate in normal scenarios, the average success rate under Gaussian noise and color jitter disturbances, and its reduction ratio compared to the normal condition.

abruptly decays to 0. This demonstrates our method's ability to adaptively respond to visual disturbances while ensuring safe and successful task execution.

Table IV reports the success rates under more challenging generalization scenarios, including Camera Shake and Unseen Block settings. To further demonstrate our method's robustness under unpredictable visual disturbances, we introduce occlusions at random timesteps and spatial locations across 50 trials, and measure their impact on task success rate. Our method demonstrates robust performance across various environmental disturbances, maintaining consistently high success rates.

C. Ablations Study

1) Importance of Uncertainty Estimate: To further verify the effectiveness of uncertainty estimation, we conduct three additional sets of experiments, as shown in Table I. First, we incorporate uncertainty estimation into the ACT method to guide action generation. This modification shows minimal improvement in the transfer (scripted) tasks, but increasing success rates for the other three tasks. Second, we introduce an alternative uncertainty estimation method, RND [21], into the ACT method, which also leads to improved performance. This demonstrates that incorporating uncertainty into the action generation process contributes to improved performance, regardless of the specific method used for uncertainty estimation. Then, we remove the uncertainty estimation component of our method and only retain the action weighting mechanism like ACT, the overall success rate decreases slightly but still exceeds that of ACT and its uncertain variant. This result highlights that the VIB-based method has significant information extraction and noise tolerance capabilities. Finally, to further verify the efficacy of our uncertainty-based framework, we substitute the uncertainty estimation module of VIB with that of RND. The resulting performance remain close to that of the original method, demonstrating the validity of the uncertainty-based strategy.

2) The Robustness Analysis: The IL methods inevitably experience performance degradation when generalizing to new scenarios with external disturbances. As shown in Fig. 4, the original ACT model exhibits a significant performance drop of up to nearly 22.0% on average under such conditions. In contrast, our approach demonstrates remarkable stability, with success rates decline by about 5.8% on average across all tested disturbance settings. This indicates that our method maintains a high level of task completion accuracy even in the presence of unforeseen perturbations, reflecting strong generalization capabilities and robustness.

3) The Influence of Lagrange Multiplier β: In the VIB framework, the Lagrange multiplier serves as a hyperparameter that balances the trade-off between retaining essential information and discarding redundant details, improving the model's generalization ability and performance stability Table II shows the effects of different Lagrange multiplier β values on the success rate. As shown in the table, the success rates generally increase as β decreases. For the transfer (scripted) task, there is little difference in success rates between $\beta = 0.001$ and $\beta = 0.0001$. However, for the other three tasks, reducing β further enhances the success rates. This suggests that choosing a smaller β value helps improve task success.

4) Multiple State Tokens: There are several reasons for using multiple state tokens as the output of the encoder network: First, in the process of estimating uncertainty, the encoder's output involves uncertainty. Using the average of multiple state tokens as an ensemble method results in more accurate and reliable uncertainty estimation. Second, similar to multi-channel processing in images or multi-head attention mechanisms, multiple tokens enable richer and more discriminative feature representation. As shown in Table V, increasing the token length improves the success rate on the insertion (scripted) task. When the length is 16 and 32, the success rates are close, with the highest overall success rate achieved at a token length of 32, reaching 40.33%. This suggests that employing multiple state tokens as the encoder network output may significantly improve task success rates. However, it's worth noting that excessively long token length (such as 64) may instead degrade the overall success rate. This implies that a trade-off is required when determining the

TABLE V
INFLUENCE OF THE MULTIPLE STATE TOKENS' LENGTH

	Insertion (Scripted)			
	Normal	Gaussian	Color	Average
$length = 1$	36	28	29	31.00
$length = 8$	38	32	35	35.00
$length = 16$	**42**	35	**42**	39.67
$length = 32$	**42**	**40**	39	**40.33**
$length = 64$	35	33	39	35.67

optimal length of multiple state tokens, balancing the benefits of capturing substantial feature information against the potential negative impacts of excessively long tokens. In conclusion, using multiple state tokens proves essential for improving the model's robustness and accuracy.

V. CONCLUSION

This study proposes VIB-ILCtrl, a novel method that integrates Variational Information Bottleneck (VIB) into imitation learning for robotic manipulation, aiming to enhance control performance through effective uncertainty estimation. The proposed framework leverages VIB to compress learned features into robust representations that generalize to unseen scenarios while simultaneously estimating the associated uncertainties. VIB-ILCtrl can accurately estimate environmental uncertainty, enabling safer operation through a threshold-based mechanism that triggers safety actions when uncertainty is high. Experimental results demonstrate that incorporating uncertainty estimation significantly improves performance by preferring reliable previous actions over potentially unreliable recent ones. Future research will explore extending this framework to more complex tasks and further refining uncertainty-guided action generation.

REFERENCES

[1] B. Wu, F. Xu, Z. He, A. Gupta, and P. K. Allen, "SQUIRL: Robust and efficient learning from video demonstration of long-horizon robotic manipulation tasks," in *Proc. 2020 IEEERSJ Int. Conf. Intell. Robots Syst.*, 2020, pp. 9720–9727.

[2] Z. Y. Ding, J. Y. Loo, V. M. Baskaran, S. G. Nurzaman, and C. P. Tan, "Predictive uncertainty estimation using deep learning for soft robot multimodal sensing," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 951–957, Apr. 2021.

[3] X. Wang et al., "Deep reinforcement learning: A survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 4, pp. 5064–5078, Apr. 2024.

[4] H. Kim, Y. Ohmura, A. Nagakubo, and Y. Kuniyoshi, "Training robots without robots: Deep imitation learning for master-to-robot policy transfer," *IEEE Robot. Autom. Lett.*, vol. 8, no. 5, pp. 2906–2913, May 2023.

[5] Z. Liu, Z. Zhang, W. Lin, X. Yu, C. Buccella, and C. Cecati, "Disturbance rejection-based motion control of linear motors via integral sliding mode observer," *IEEE Trans. Ind. Electron.*, vol. 72, no. 3, pp. 3061–3071, Mar. 2025.

[6] B. Zheng, S. Verma, J. Zhou, I. W. Tsang, and F. Chen, "Imitation learning: Progress, taxonomies and challenges," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 5, pp. 6322–6337, May 2024.

[7] R. Pérez-Dattari and J. Kober, "Stable motion primitives via imitation and contrastive learning," *IEEE Trans. Robot.*, vol. 39, no. 5, pp. 3909–3928, Oct. 2023.

[8] M. Xu, Z. Xu, C. Chi, M. Veloso, and S. Song, "XSkill: Cross embodiment skill discovery," in *Proc. Conf. Robot Learn.*, 2023, pp. 3536–3555.

[9] N. Gavenski, O. Rodrigues, and M. Luck, "Imitation learning: A survey of learning methods, environments and metrics," 2024, *arXiv:2404.19456*.

[10] E. Rosete-Beas, O. Mees, G. Kalweit, J. Boedecker, and W. Burgard, "Latent plans for task-agnostic offline reinforcement learning," in *Proc. Conf. Robot Learn.*, 2022, pp. 1–12.

[11] F. D. Felice, S. D'Avella, A. Remus, P. Tripicchio, and C. A. Avizzano, "One-shot imitation learning with graph neural networks for pick-and-place manipulation tasks," *IEEE Robot. Autom. Lett.*, vol. 8, no. 9, pp. 5926–5933, Sep. 2023.

[12] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning fine-grained bimanual manipulation with low-cost hardware," *Robot. Sci. Syst.*, 2023.

[13] B. D. Ziebart, J. A. Bagnell, and A. K. Dey, "Modeling interaction via the principle of maximum causal entropy," in *Proc. 27th Int. Conf. Int. Conf. Mach. Learn.*, Madison, WI, USA, 2010, pp. 1255–1262.

[14] X. B. Peng, A. Kanazawa, S. Toyer, P. Abbeel, and S. Levine, "Variational discriminator bottleneck: Improving imitation learning, inverse RL, and GANs by constraining information flow," in *Proc. Int. Conf. Learn. Representations*, Apr. 2014, pp. 1–14.

[15] W. Sun, A. Vemula, B. Boots, and D. Bagnell, "Provably efficient imitation learning from observation alone," in *Proc. 36th Int. Conf. Mach. Learn.*, 2019, pp. 6036–6045.

[16] A. Kendall and Y. Gal, "What uncertainties do we need in Bayesian deep learning for computer vision?," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, vol. 30, pp. 1–11.

[17] PujolOriol MenaJosé and VitriàJordi, "A survey on uncertainty estimation in deep learning classification systems from a Bayesian perspective," *ACM Comput. Surv.*, vol. 54, pp. 1–35, Oct. 2021.

[18] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 6405–6416.

[19] G. An, S. Moon, J.-H. Kim, and H. O. Song, "Uncertainty-based offline reinforcement learning with diversified Q-ensemble," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, vol. 34, pp. 7436–7447.

[20] S. K. S. Ghasemipour, S. S. Gu, and O. Nachum, "Why so pessimistic? Estimating uncertainties for offline RL through ensembles, and why their independence matters," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 1–13.

[21] Y. Burda, H. Edwards, A. Storkey, and O. Klimov, "Exploration by random network distillation," in *Proc. Int. Conf. Learn. Representations*, Apr. 2014, pp. 1–13.

[22] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," 2000, *arXiv:Physics/0004057*.

[23] A. A. Alemi, I. Fischer, J. V. Dillon, and K. Murphy, "Deep variational information bottleneck," in *Proc. Int. Conf. Learn. Represention*, Feb. 2017, pp. 1–16.

[24] A. A. Alemi, I. Fischer, and J. V. Dillon, "Uncertainty in the variational information bottleneck," 2018, *arXiv:1807.00906*.

[25] K. Hsu, M. J. Kim, R. Rafailov, J. Wu, and C. Finn, "Vision-based manipulators need to also see from their hands," in *Proc. Int. Conf. Learn. Representions*, 2021, pp. 1–13.

[26] Y. Du et al., "Learning to learn with variational information bottleneck for domain generalization," in *Proc. IEEE Eur. Conf. Comput. Vis.*, 2020, pp. 200–216.

[27] S. Kudo, N. Ono, S. Kanaya, and M. Huang, "Flexible variational information bottleneck: Achieving diverse compression with a single training," *Neurocomputing*, vol. 640, Aug. 2025, Art. no. 130198.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. 2016 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[29] J. Li, D. Li, S. Savarese, and S. Hoi, "BLIP-2: Bootstrapping language-image pre-training with frozen image encoders and large language models," in *Proc. 40th Int. Conf. Mach. Learn.*, 2023, pp. 19730–19742.

[30] C. Chi et al., "Diffusion policy: Visuomotor policy learning via action diffusion," *Robot. Sci. Syst.*, 2023.

[31] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. Int. Conf. Learn. Representations*, Apr. 2014, pp. 1–9.

[32] T. Zhang et al., "Deep imitation learning for complex manipulation tasks from virtual reality teleoperation," in *Proc. 2018 IEEE Int. Conf. Robot. Autom.*, Brisbane, QLD, Australia, 2018, pp. 5628–5635.