Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Bonn-Aachen International Center for Information Technology

Institute for AI and Autonomous Systems

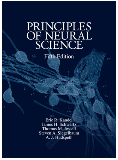
Visuomotor Policies

From Biology to Robots

Dr. Alex Mitrevski
Master of Autonomous Systems

Structure

Master's Thesis
Visuomotor Policy Learning for Predictive Manipulation
Anirudh Narasimamurthy Jayasimha



- Visuomotor policy preliminaries
- Biological insights
- Visuomotor robot policies

Visuomotor Policy Preliminaries

What is a Visuomotor Policy?

- As discussed in our introductory lecture, **a visuomotor policy uses visual information to guide the production of motor actions**

What is a Visuomotor Policy?

- As discussed in our introductory lecture, **a visuomotor policy uses visual information to guide the production of motor actions**
- This means that a visuomotor policy creates **a direct functional mapping between visual observations and the corresponding actions**

What is a Visuomotor Policy?

- As discussed in our introductory lecture, **a visuomotor policy uses visual information to guide the production of motor actions**
- This means that a visuomotor policy creates **a direct functional mapping between visual observations and the corresponding actions**
- Typically, a visuomotor policy is not based on visual information only, but **visual observations are combined with other sensory modalities**

Why Visuomotor Policies?

- Visual information is perhaps the richest source of information that a robot can consume while acting — **visuomotor policies enable the acting process to benefit from such information**

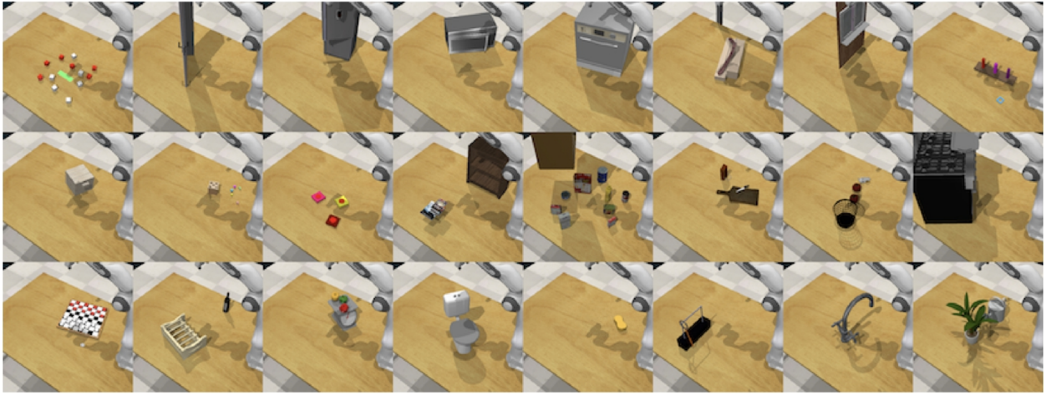
Why Visuomotor Policies?

- Visual information is perhaps the richest source of information that a robot can consume while acting — **visuomotor policies enable the acting process to benefit from such information**
- For many robotics tasks, task completion depends on features that can be best observed in visual data (e.g. pouring a liquid to a desired level) — **processing visual information can thus enable functionalities that are otherwise difficult or impossible to achieve**

Why Visuomotor Policies?

- Visual information is perhaps the richest source of information that a robot can consume while acting — **visuomotor policies enable the acting process to benefit from such information**
- For many robotics tasks, task completion depends on features that can be best observed in visual data (e.g. pouring a liquid to a desired level) — **processing visual information can thus enable functionalities that are otherwise difficult or impossible to achieve**
- Visual information is also essential to consider in scenarios where a robot should be able to adapt its behaviour based on the observations of a cooperating human — **visuomotor policies can enable more effective human-robot collaboration**

Visual Information Can Be Beneficial for Many Robotics Scenarios



S. James et al., “RLBench: The Robot Learning Benchmark & Learning Environment,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3019–3026, Apr. 2020.

Visuomotor Policies vs. Visual Servoing

- Visuomotor policies are **conceptually equivalent to visual servoing control**, where visual information is used as a control feedback signal

Visuomotor Policies vs. Visual Servoing

- Visuomotor policies are **conceptually equivalent to visual servoing control**, where visual information is used as a control feedback signal
- The sense in which we use the term visuomotor policy in this class, and in the contemporary literature on robot learning, is that of **a policy that is learned by a robot**

Visuomotor Policies vs. Visual Servoing

- Visuomotor policies are **conceptually equivalent to visual servoing control**, where visual information is used as a control feedback signal
- The sense in which we use the term visuomotor policy in this class, and in the contemporary literature on robot learning, is that of **a policy that is learned by a robot**
- Traditional visual servoing, on the other hand, **combines control theory with traditional computer vision**, such that the error estimation is done through the computation of visual features

Visuomotor Policies vs. Visual Servoing

- Visuomotor policies are **conceptually equivalent to visual servoing control**, where visual information is used as a control feedback signal
- The sense in which we use the term visuomotor policy in this class, and in the contemporary literature on robot learning, is that of **a policy that is learned by a robot**
- Traditional visual servoing, on the other hand, **combines control theory with traditional computer vision**, such that the error estimation is done through the computation of visual features
- **The learning aspect is thus the primary distinction** between (traditional) visual servoing and the (learning-based) visuomotor policies that we discuss in this lecture

Biological Insights

Visuomotor Policies in Biological Systems

- Visuomotor policies as a concept are **directly inspired from the operation of biological systems**
  - Particularly humans exhibit an exceptional ability to use sensory information for performing a variety of coarse and fine motions

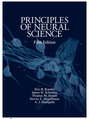
Visuomotor Policies in Biological Systems

- Visuomotor policies as a concept are **directly inspired from the operation of biological systems**
  - Particularly humans exhibit an exceptional ability to use sensory information for performing a variety of coarse and fine motions
- Considering visuomotor policies in biological systems is useful because it can inform the **design and development of visuomotor robot policies** and the **components that make such policies work in practice**

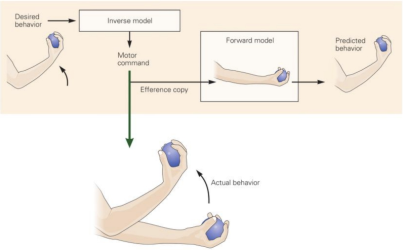
Visuomotor Policies in Biological Systems

- Visuomotor policies as a concept are **directly inspired from the operation of biological systems**
  - Particularly humans exhibit an exceptional ability to use sensory information for performing a variety of coarse and fine motions
- Considering visuomotor policies in biological systems is useful because it can inform the **design and development of visuomotor robot policies** and the **components that make such policies work in practice**
- A study of biological visuomotor policies can also show **how contemporary approaches to visuomotor robot policies compare to their biological counterparts** and **what aspects are missing for obtaining policies that are as versatile**

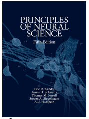
Motion Simulation Using Forward Models

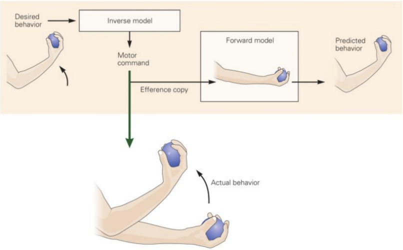


- According to studies of the central nervous system, the brain stores various **internal models** that are used during motor movements



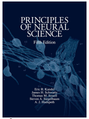
Motion Simulation Using Forward Models

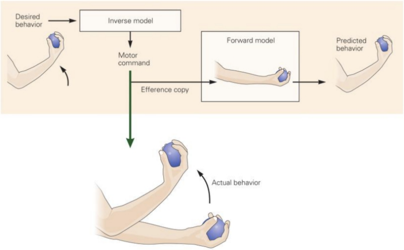




- According to studies of the central nervous system, the brain stores various **internal models** that are used during motor movements
- In addition to models that are responsible for generating actual motor commands, the central nervous system likely has a **forward model that can predict the expected sensor measurement resulting from a given motor command**

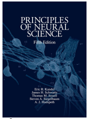
Motion Simulation Using Forward Models

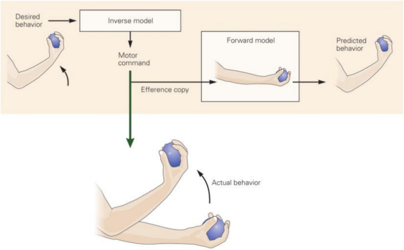




- According to studies of the central nervous system, the brain stores various **internal models** that are used during motor movements
- In addition to models that are responsible for generating actual motor commands, the central nervous system likely has a **forward model that can predict the expected sensor measurement resulting from a given motor command**
- An **efferent signal** is a motion signal that is sent by the central nervous system; such a signal sent to the forward model is called an **efferent copy**

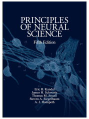
Motion Simulation Using Forward Models

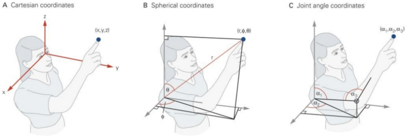




- According to studies of the central nervous system, the brain stores various **internal models** that are used during motor movements
- In addition to models that are responsible for generating actual motor commands, the central nervous system likely has a **forward model that can predict the expected sensor measurement resulting from a given motor command**
- An **efferent signal** is a motion signal that is sent by the central nervous system; such a signal sent to the forward model is called an **efferent copy**
- Forward models, and the associated **inverse models that generate motor commands**, are learned during early childhood movements and improved continually

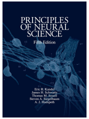
Movement Coordinate Systems

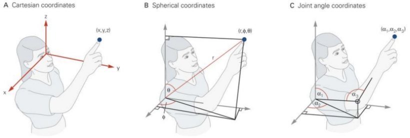




- For generating motion, the brain can use various **intrinsic or extrinsic coordinate systems**
  - According to studies of neuron firing patterns, **different coordinate systems are encoded in different parts of the brain**
  - The used coordinate system is **task-dependent**

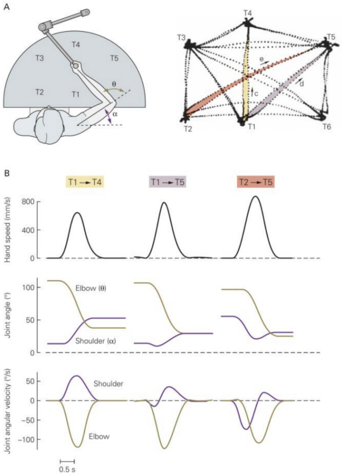
Movement Coordinate Systems

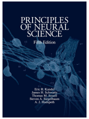




- For generating motion, the brain can use various **intrinsic or extrinsic coordinate systems**
  - According to studies of neuron firing patterns, **different coordinate systems are encoded in different parts of the brain**
  - The used coordinate system is **task-dependent**
- The coordinate system used for a task is typically not known directly, but can be determined:
  - **by studying neuron firing patterns during a task** or
  - **by analysing motion errors with respect to different variables and coordinate systems**

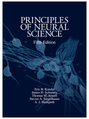
Stereotypical Motions and Motion Primitives

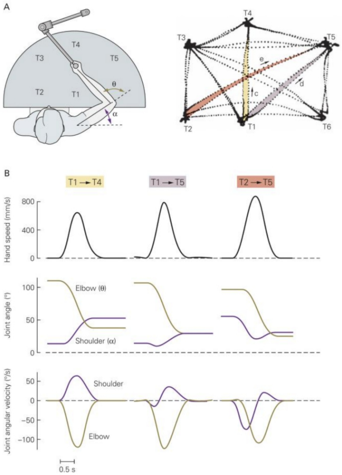




- Studies of reaching motions have shown that **motions usually exhibit stereotypical patterns**
    - Such studies particularly show that **straight-line motions are usually preferred to more complex arc motions**, suggesting that motion trajectories are typically planned with respect to the hand

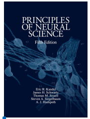
Stereotypical Motions and Motion Primitives

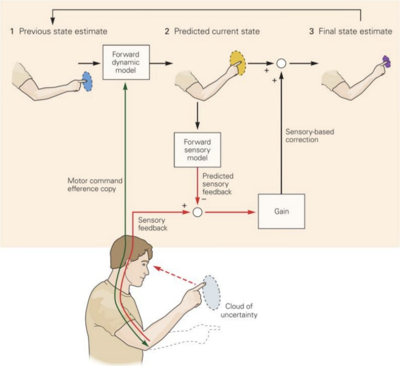




- Studies of reaching motions have shown that **motions usually exhibit stereotypical patterns**
  - Such studies particularly show that **straight-line motions are usually preferred to more complex arc motions**, suggesting that motion trajectories are typically planned with respect to the hand
- Movement studies also suggest that **complex motions are combinations of well-defined motion primitives**
  - The choice of primitives to complex motions is thought to be done **by optimising a cost associated with each primitive**
  - The applied cost function may **depend on both the task and on the properties of motion**
  - Both evolution and learning likely have an effect on the primitives and the used cost functions

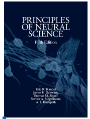
Motion Using Feedforward and Feedback Control

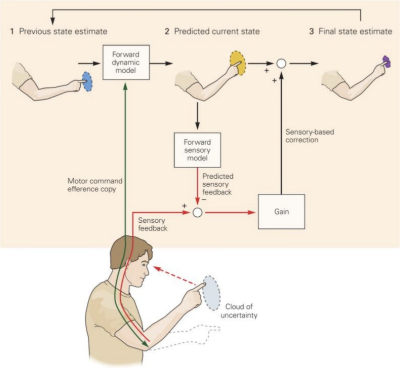




- Processing sensory feedback **introduces motion delays**
  - For instance, processing and sending visual signals has been shown to introduce a $200ms$ delay

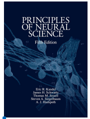
Motion Using Feedforward and Feedback Control

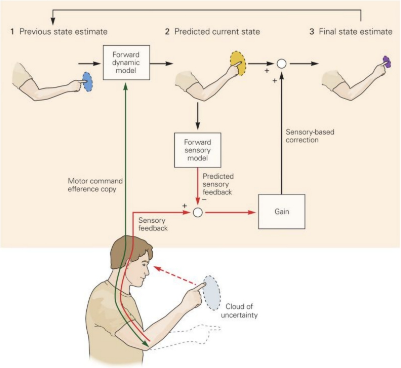




- Processing sensory feedback **introduces motion delays**
  - For instance, processing and sending visual signals has been shown to introduce a $200ms$ delay
- For this reason, **motions are sometimes performed using feedforward control**
  - Such motions have particularly been observed during the initial stages of motion

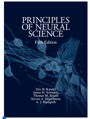
Motion Using Feedforward and Feedback Control

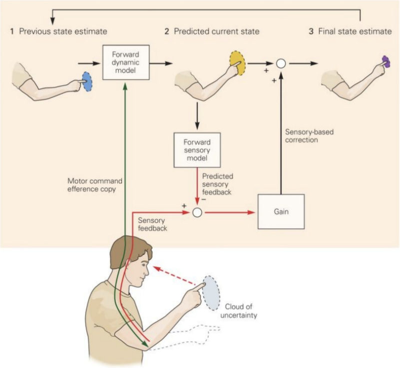




- Processing sensory feedback **introduces motion delays**
  - For instance, processing and sending visual signals has been shown to introduce a $200ms$ delay
- For this reason, **motions are sometimes performed using feedforward control**
  - Such motions have particularly been observed during the initial stages of motion
- Pure feedforward control would require too accurate inverse models; **most motions are typically performed by following-up feedforward with feedback control**
  - Motion inaccuracies increase with increased motion speed

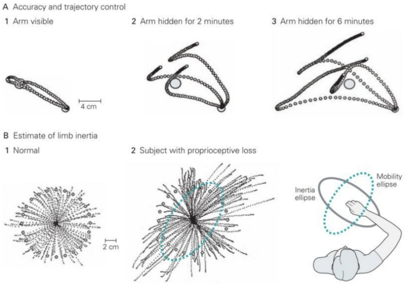
Motion Using Feedforward and Feedback Control





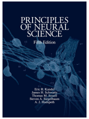
- Processing sensory feedback **introduces motion delays**
  - For instance, processing and sending visual signals has been shown to introduce a $200ms$ delay
- For this reason, **motions are sometimes performed using feedforward control**
  - Such motions have particularly been observed during the initial stages of motion
- Pure feedforward control would require too accurate inverse models; **most motions are typically performed by following-up feedforward with feedback control**
  - Motion inaccuracies increase with increased motion speed
- To account for delays in sensory feedback and inaccuracies in pure motion prediction, **an observer model** can be employed

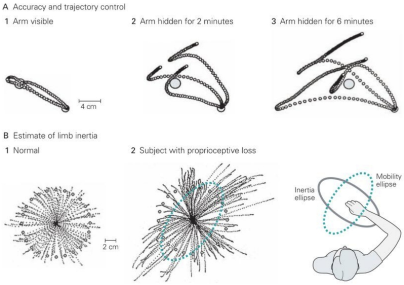
Importance of Proprioception During Motion



- In studies of people who have lost the ability to make proprioceptive measurements, it has been shown that **proprioception is essential for moving accurately**

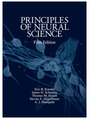
Importance of Proprioception During Motion

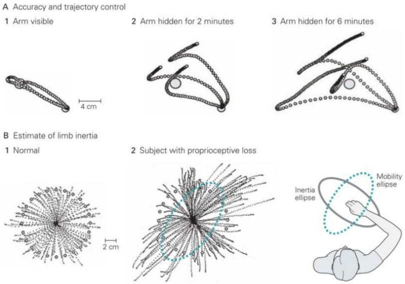




- In studies of people who have lost the ability to make proprioceptive measurements, it has been shown that **proprioception is essential for moving accurately**
- The loss of proprioception can be **compensated by visual feedback** received by seeing the moved limbs

Importance of Proprioception During Motion

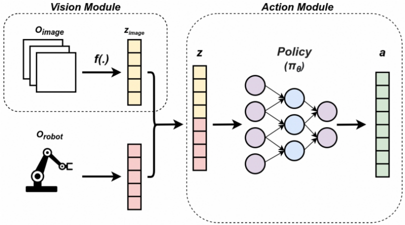




- In studies of people who have lost the ability to make proprioceptive measurements, it has been shown that **proprioception is essential for moving accurately**
- The loss of proprioception can be **compensated by visual feedback** received by seeing the moved limbs
- This overreliance on visual information means that **acting without visual information becomes impossible**

Visuomotor Robot Policies

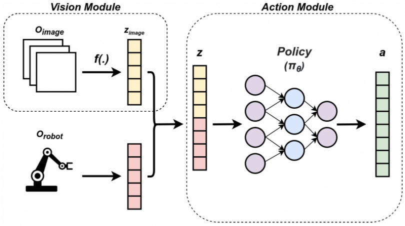
Basic Visuomotor Policy



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- A typical visuomotor robot policy combines at least two types of sensor data:
  - **exteroceptive visual data**
  - **proprioceptive robot data**, typically in the form of joint positions

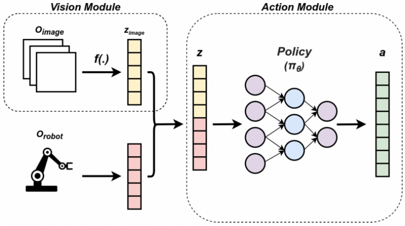
Basic Visuomotor Policy



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- A typical visuomotor robot policy combines at least two types of sensor data:
  - **exteroceptive visual data**
  - **proprioceptive robot data**, typically in the form of joint positions
- A visuomotor policy **fuses these data** in order to produce a desired policy output

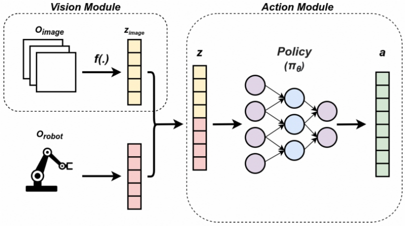
Basic Visuomotor Policy



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- A typical visuomotor robot policy combines at least two types of sensor data:
  - **exteroceptive visual data**
  - **proprioceptive robot data**, typically in the form of joint positions
- A visuomotor policy **fuses these data** in order to produce a desired policy output
- Visuomotor policies are predominantly represented by neural networks
  - Here, fusion is typically done at the level of **latent representation**, which are separately extracted from the individual modalities

Basic Visuomotor Policy

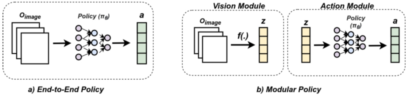


A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- A typical visuomotor robot policy combines at least two types of sensor data:
  - **exteroceptive visual data**
  - **proprioceptive robot data**, typically in the form of joint positions
- A visuomotor policy **fuses these data** in order to produce a desired policy output
- Visuomotor policies are predominantly represented by neural networks
  - Here, fusion is typically done at the level of **latent representation**, which are separately extracted from the individual modalities
- The visual processing module may process **a single image**, **a series of images**, or encode visual state memory through a **recurrent representation**

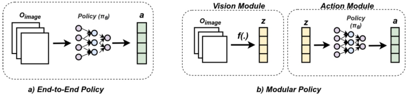
Feature Extraction from Visual Data

- One dilemma when designing a visuomotor policy is whether to use a modular or an end-to-end policy



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

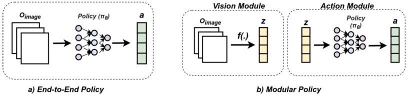
Feature Extraction from Visual Data



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- One dilemma when designing a visuomotor policy is whether to use a modular or an end-to-end policy
- An end-to-end policy **learns a feature extractor jointly with learning a policy**, while **a modular policy uses features from a pretrained visual feature extractor**

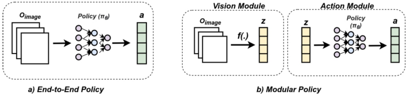
Feature Extraction from Visual Data



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- One dilemma when designing a visuomotor policy is whether to use a modular or an end-to-end policy
- An end-to-end policy **learns a feature extractor jointly with learning a policy**, while **a modular policy uses features from a pretrained visual feature extractor**
- A modular design can **speed up the policy learning process** — fewer parameters need to be optimised during policy learning and so learning requires fewer iterations — but **does not allow the visual component to adapt itself to the task of interest**

Feature Extraction from Visual Data

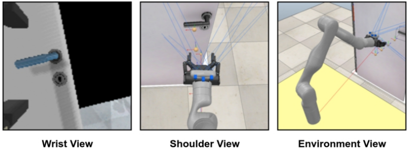


A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- One dilemma when designing a visuomotor policy is whether to use a modular or an end-to-end policy
- An end-to-end policy **learns a feature extractor jointly with learning a policy**, while **a modular policy uses features from a pretrained visual feature extractor**
- A modular design can **speed up the policy learning process** — fewer parameters need to be optimised during policy learning and so learning requires fewer iterations — but **does not allow the visual component to adapt itself to the task of interest**
- **End-to-end learning can also be performed with a pretrained visual component** — here, the visual component is **trained on an auxiliary task** and then **further optimised together with the policy**

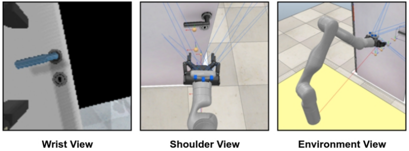
Camera Positioning for Visuomotor Control

- Visuomotor policies can be trained with cameras placed at different positions — this is typically a task-dependent aspect



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

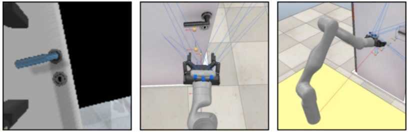
Camera Positioning for Visuomotor Control



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- Visuomotor policies can be trained with cameras placed at different positions — this is typically a task-dependent aspect
- Different camera positions have several implications for visuomotor policies:
  - Environment cameras **can be placed flexibly so that the robot has the best view of the relevant parts of a scene**, but **require a high level of environment engineering** — not generally applicable

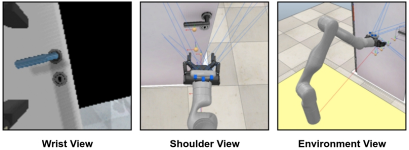
Camera Positioning for Visuomotor Control



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- Visuomotor policies can be trained with cameras placed at different positions — this is typically a task-dependent aspect
- Different camera positions have several implications for visuomotor policies:
  - Environment cameras **can be placed flexibly so that the robot has the best view of the relevant parts of a scene**, but **require a high level of environment engineering** — not generally applicable
  - Shoulder cameras **generally provide a solid view of the frontal scene**, but **a robot needs to deal with self-occlusions** caused by the motion of the manipulator

Camera Positioning for Visuomotor Control



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- Visuomotor policies can be trained with cameras placed at different positions — this is typically a task-dependent aspect
- Different camera positions have several implications for visuomotor policies:
  - Environment cameras **can be placed flexibly so that the robot has the best view of the relevant parts of a scene**, but **require a high level of environment engineering** — not generally applicable
  - Shoulder cameras **generally provide a solid view of the frontal scene**, but **a robot needs to deal with self-occlusions** caused by the motion of the manipulator
  - Wrist cameras **enable a robot to have a close-up perspective of the scene during execution**, but **typically provide a small scene view** — mostly applicable during the final segments of a task

Training Data Sources for Visuomotor Policies

- Due to the fact that they process raw, unstructured state information, **visuomotor policies are typically difficult to train and require many learning iterations**

Training Data Sources for Visuomotor Policies

- Due to the fact that they process raw, unstructured state information, **visuomotor policies are typically difficult to train and require many learning iterations**
- The question of the origin of data used for visuomotor policy learning is thus an important one, as **the source can have a significant effect on the training efficiency**

Training Data Sources for Visuomotor Policies

- Due to the fact that they process raw, unstructured state information, **visuomotor policies are typically difficult to train and require many learning iterations**
- The question of the origin of data used for visuomotor policy learning is thus an important one, as **the source can have a significant effect on the training efficiency**
- There are multiple data sources that can be utilised during training:
  - Data can be collected **directly during policy learning**, which would require many interactions with the environment

Training Data Sources for Visuomotor Policies

- Due to the fact that they process raw, unstructured state information, **visuomotor policies are typically difficult to train and require many learning iterations**
- The question of the origin of data used for visuomotor policy learning is thus an important one, as **the source can have a significant effect on the training efficiency**
- There are multiple data sources that can be utilised during training:
  - Data can be collected **directly during policy learning**, which would require many interactions with the environment
  - As mentioned before, data for solving **an auxiliary task** can be collected prior to policy learning; these data can be used to initialise the feature extractor

Training Data Sources for Visuomotor Policies

- Due to the fact that they process raw, unstructured state information, **visuomotor policies are typically difficult to train and require many learning iterations**
- The question of the origin of data used for visuomotor policy learning is thus an important one, as **the source can have a significant effect on the training efficiency**
- There are multiple data sources that can be utilised during training:
  - Data can be collected **directly during policy learning**, which would require many interactions with the environment
  - As mentioned before, data for solving **an auxiliary task** can be collected prior to policy learning; these data can be used to initialise the feature extractor
  - The feature extractors or the complete policy can be **initialised from demonstrations**

Training Data Sources for Visuomotor Policies

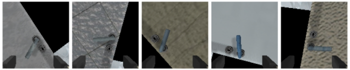
- Due to the fact that they process raw, unstructured state information, **visuomotor policies are typically difficult to train and require many learning iterations**
- The question of the origin of data used for visuomotor policy learning is thus an important one, as **the source can have a significant effect on the training efficiency**
- There are multiple data sources that can be utilised during training:
  - Data can be collected **directly during policy learning**, which would require many interactions with the environment
  - As mentioned before, data for solving **an auxiliary task** can be collected prior to policy learning; these data can be used to initialise the feature extractor
  - The feature extractors or the complete policy can be **initialised from demonstrations**
  - A policy can be trained using data collected **in simulation** and then potentially **fine-tuned using real-world data**

Training Data Sources for Visuomotor Policies

- Due to the fact that they process raw, unstructured state information, **visuomotor policies are typically difficult to train and require many learning iterations**
- The question of the origin of data used for visuomotor policy learning is thus an important one, as **the source can have a significant effect on the training efficiency**
- There are multiple data sources that can be utilised during training:
  - Data can be collected **directly during policy learning**, which would require many interactions with the environment
  - As mentioned before, data for solving **an auxiliary task** can be collected prior to policy learning; these data can be used to initialise the feature extractor
  - The feature extractors or the complete policy can be **initialised from demonstrations**
  - A policy can be trained using data collected **in simulation** and then potentially **fine-tuned using real-world data**
- **The source of data used for policy learning is often task-dependent**; for instance, demonstrations are typically a valuable source of expert data, but can be difficult or time-consuming to collect

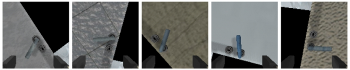
Domain Randomisation

- The neural networks that are incorporated into policies are usually based on convolutional neural networks; these are **invariant to translations by design**, but not to other properties, such as brightness or background variations



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

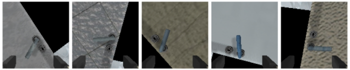
Domain Randomisation



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- The neural networks that are incorporated into policies are usually based on convolutional neural networks; these are **invariant to translations by design**, but not to other properties, such as brightness or background variations
- To make a policy invariant to such properties, **domain randomisation** should be done during visual data collection
  - Domain randomisation virtually increases the visual training set size by introducing images that are modified versions of the collected images

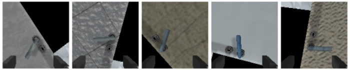
Domain Randomisation



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- The neural networks that are incorporated into policies are usually based on convolutional neural networks; these are **invariant to translations by design**, but not to other properties, such as brightness or background variations
- To make a policy invariant to such properties, **domain randomisation** should be done during visual data collection
  - Domain randomisation virtually increases the visual training set size by introducing images that are modified versions of the collected images
- Typical domain randomisation operations include **image flips, brightness and contrast changes, as well as pasting objects onto different backgrounds**

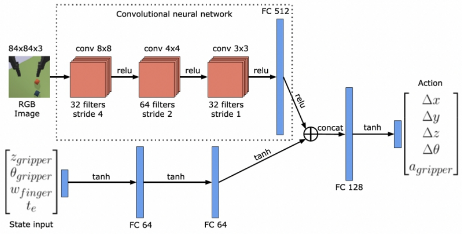
Domain Randomisation



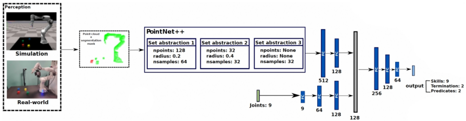
A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- The neural networks that are incorporated into policies are usually based on convolutional neural networks; these are **invariant to translations by design**, but not to other properties, such as brightness or background variations
- To make a policy invariant to such properties, **domain randomisation** should be done during visual data collection
  - Domain randomisation virtually increases the visual training set size by introducing images that are modified versions of the collected images
- Typical domain randomisation operations include **image flips, brightness and contrast changes, as well as pasting objects onto different backgrounds**
- Randomisation during training has traditionally been the easiest to perform in simulations, but generative image models can also produce useful, photorealistic augmented data

Sensory Input Varieties



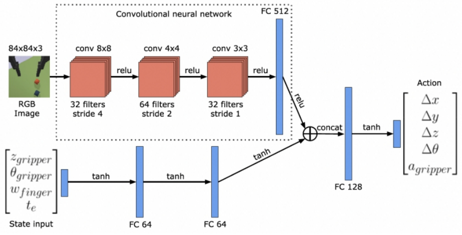
L. Hermann et al., "Adaptive Curriculum Generation from Demonstrations for Sim-to-Real Visuomotor Control," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020, pp. 6498–6505.



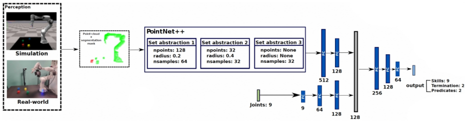
S. Mukherjee et al., "Reactive Long Horizon Task Execution via Visual Skill and Precondition Models," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2021, pp. 5717–5724.

- The visual data used as an input to a visuomotor policy can take different shapes; this has an effect on the design of the network architecture used for visual data processing

Sensory Input Varieties



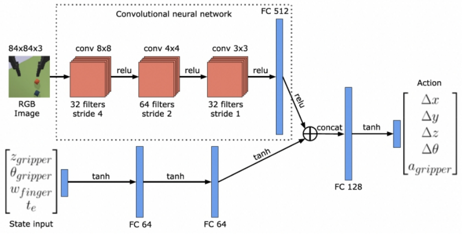
L. Hermann et al., "Adaptive Curriculum Generation from Demonstrations for Sim-to-Real Visuomotor Control," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020, pp. 6498–6505.



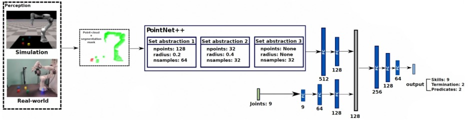
S. Mukherjee et al., "Reactive Long Horizon Task Execution via Visual Skill and Precondition Models," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2021, pp. 5717–5724.

- The visual data used as an input to a visuomotor policy can take different shapes; this has an effect on the design of the network architecture used for visual data processing
- The most common format is that of **a raw RGB image** that is typically processed by **a convolutional neural network**; the output of a fully connected layer is then used as a visual feature representation

Sensory Input Varieties



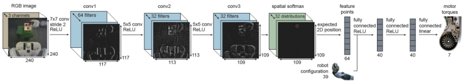
L. Hermann et al., "Adaptive Curriculum Generation from Demonstrations for Sim-to-Real Visuomotor Control," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020, pp. 6498–6505.



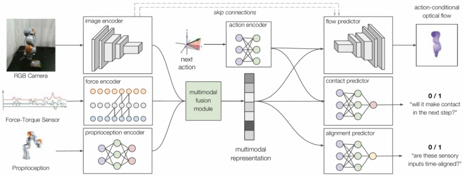
S. Mukherjee et al., "Reactive Long Horizon Task Execution via Visual Skill and Precondition Models," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2021, pp. 5717–5724.

- The visual data used as an input to a visuomotor policy can take different shapes; this has an effect on the design of the network architecture used for visual data processing
- The most common format is that of **a raw RGB image** that is typically processed by **a convolutional neural network**; the output of a fully connected layer is then used as a visual feature representation
- Alternative representations are naturally possible as well and may work better for point cloud inputs, for instance **segmentation masks**

Multimodality



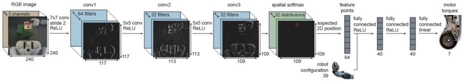
S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-End Training of Deep Visuomotor Policies," *Journal Machine Learning Research*, vol. 17, no. 1, pp. 1334–1373, Jan. 2016.



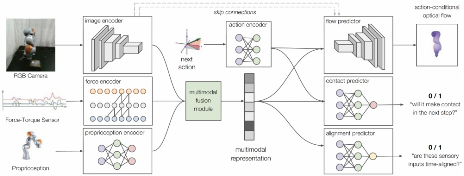
M. A. Lee et al., "Making Sense of Vision and Touch: Self-Supervised Learning of Multimodal Representations for Contact-Rich Tasks," in *Proc. Int. Conf. Robotics and Automation (ICRA)*, 2019, pp. 8943–8950.

- Multimodal policy network have **multiple modality-specific branches** that are then fused to produce a joint feature vector that represents the input to a policy

Multimodality



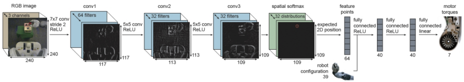
S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-End Training of Deep Visuomotor Policies," *Journal Machine Learning Research*, vol. 17, no. 1, pp. 1334–1373, Jan. 2016.



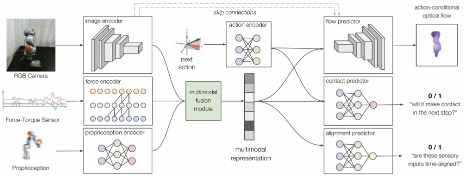
M. A. Lee et al., "Making Sense of Vision and Touch: Self-Supervised Learning of Multimodal Representations for Contact-Rich Tasks," in *Proc. Int. Conf. Robotics and Automation (ICRA)*, 2019, pp. 8943–8950.

- Multimodal policy network have **multiple modality-specific branches** that are then fused to produce a joint feature vector that represents the input to a policy
- Fusing visual information with proprioceptive information about a robot's joints is one example of multimodality — **two modalities are processed individually and are then combined to form the policy input**

Multimodality



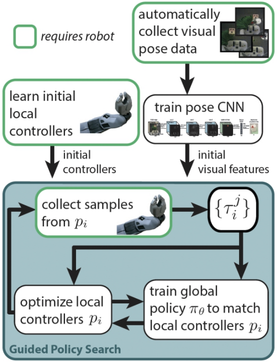
S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-End Training of Deep Visuomotor Policies," *Journal Machine Learning Research*, vol. 17, no. 1, pp. 1334–1373, Jan. 2016.



M. A. Lee et al., "Making Sense of Vision and Touch: Self-Supervised Learning of Multimodal Representations for Contact-Rich Tasks," in *Proc. Int. Conf. Robotics and Automation (ICRA)*, 2019, pp. 8943–8950.

- Multimodal policy network have **multiple modality-specific branches** that are then fused to produce a joint feature vector that represents the input to a policy
- Fusing visual information with proprioceptive information about a robot's joints is one example of multimodality — **two modalities are processed individually and are then combined to form the policy input**
- For some practical tasks, **visual and proprioceptive information may be insufficient for successfully completing a task** (e.g. lifting an object with an unknown weight)
  - Contact-heavy tasks in particular require taking into account additional information, such as force measurements or tactile feedback

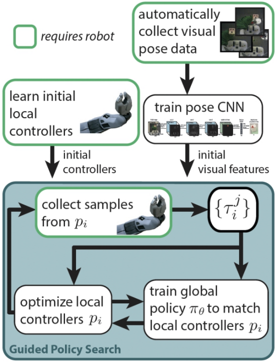
Guided Visuomotor Policy Learning



S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-End Training of Deep Visuomotor Policies," *Journal Machine Learning Research*, vol. 17, no. 1, pp. 1334–1373, Jan. 2016.

- The discussion thus far considers model-free policy learning, but **learning visuomotor policies based on models** is both possible and often desirable
  - Model-based learning is more sample-efficient than model-free learning, as it benefits from predictions made by **forward models**

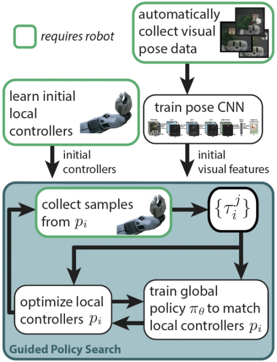
Guided Visuomotor Policy Learning



S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-End Training of Deep Visuomotor Policies," *Journal Machine Learning Research*, vol. 17, no. 1, pp. 1334–1373, Jan. 2016.

- The discussion thus far considers model-free policy learning, but **learning visuomotor policies based on models** is both possible and often desirable
  - Model-based learning is more sample-efficient than model-free learning, as it benefits from predictions made by **forward models**
- To be useful for learning, forward models **need to be initialised appropriately**, but they can also be **further optimised jointly with the policy**

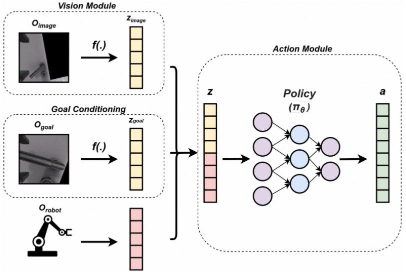
Guided Visuomotor Policy Learning



S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-End Training of Deep Visuomotor Policies," *Journal Machine Learning Research*, vol. 17, no. 1, pp. 1334–1373, Jan. 2016.

- The discussion thus far considers model-free policy learning, but **learning visuomotor policies based on models** is both possible and often desirable
  - Model-based learning is more sample-efficient than model-free learning, as it benefits from predictions made by **forward models**
- To be useful for learning, forward models **need to be initialised appropriately**, but they can also be **further optimised jointly with the policy**
- Forward models may also be **useful during policy execution**, which is a desirable property from a biological point of view (as we have seen before)

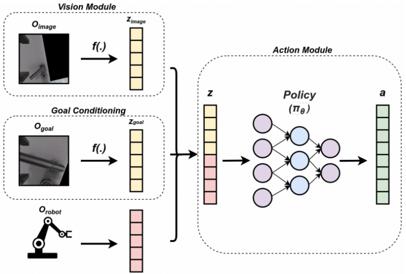
Goal-Conditioned Visuomotor Policies



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- So far, we have only seen visuomotor policies of the form $\pi : S \rightarrow A$, which do not encode information about the goal that a robot is trying to achieve

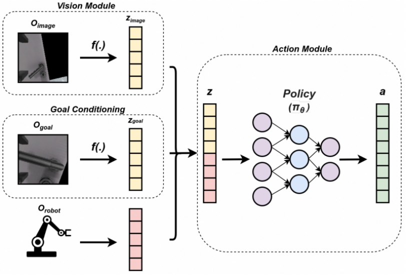
Goal-Conditioned Visuomotor Policies



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- So far, we have only seen visuomotor policies of the form $\pi : S \rightarrow A$, which do not encode information about the goal that a robot is trying to achieve
- A **goal-conditioned policy** is instead of the form $\pi : (S, C) \rightarrow A$, where $\boldsymbol{c} \in C$ represents the goal of the execution
  - The goal acts as a **context vector** that defines a task family over goals

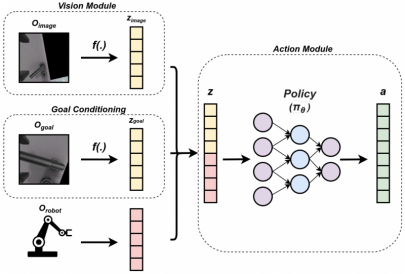
Goal-Conditioned Visuomotor Policies



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- So far, we have only seen visuomotor policies of the form $\pi : S \rightarrow A$, which do not encode information about the goal that a robot is trying to achieve
- A **goal-conditioned policy** is instead of the form $\pi : (S, C) \rightarrow A$, where $\boldsymbol{c} \in C$ represents the goal of the execution
  - The goal acts as a **context vector** that defines a task family over goals
- Goal conditioning can **stabilise the policy learning**, as goals provide guidance on the state that the robot should aim to achieve during execution

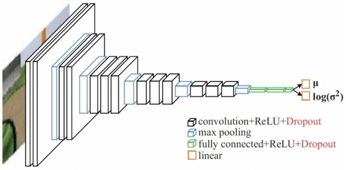
Goal-Conditioned Visuomotor Policies



A. Narasimamurthy, "Visuomotor Policy Learning for Predictive Manipulation," Master's thesis, Hochschule Bonn-Rhein-Sieg, May 2021.

- So far, we have only seen visuomotor policies of the form $\pi : S \rightarrow A$, which do not encode information about the goal that a robot is trying to achieve
- A **goal-conditioned policy** is instead of the form $\pi : (S, C) \rightarrow A$, where $\boldsymbol{c} \in C$ represents the goal of the execution
  - The goal acts as a **context vector** that defines a task family over goals
- Goal conditioning can **stabilise the policy learning**, as goals provide guidance on the state that the robot should aim to achieve during execution
- The goal in a visuomotor policy **encodes a representation of a goal image** — for instance, as a latent feature representation

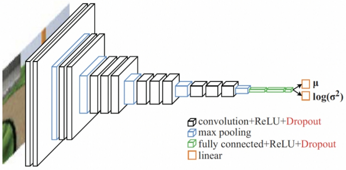
Representing Uncertainty in Policies



K. Lee, K. Saigol and E. A. Theodorou, "Early Failure Detection of Deep End-to-End Control Policy by Reinforcement Learning," in *Proc. Int. Conf. Robotics and Automation (ICRA)*, 2019, pp. 8543–8549.

- Typical visuomotor robot policies produce motor commands, but **have no notion about the quality of the selected motions under certain conditions**
  - Resulting policies may be unsafe as a result

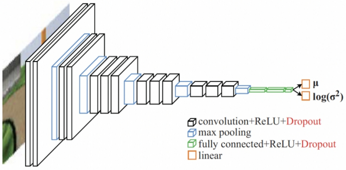
Representing Uncertainty in Policies



K. Lee, K. Saigol and E. A. Theodorou, "Early Failure Detection of Deep End-to-End Control Policy by Reinforcement Learning," in *Proc. Int. Conf. Robotics and Automation (ICRA)*, 2019, pp. 8543–8549.

- Typical visuomotor robot policies produce motor commands, but **have no notion about the quality of the selected motions under certain conditions**
  - Resulting policies may be unsafe as a result
- Representing uncertainty over the output requires **modelling the output as a distribution** rather than as a maximum likelihood point estimate — Bayesian inference is required

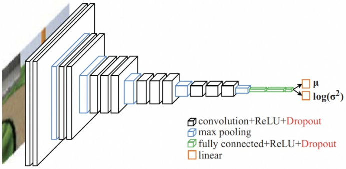
Representing Uncertainty in Policies



K. Lee, K. Saigol and E. A. Theodorou, "Early Failure Detection of Deep End-to-End Control Policy by Reinforcement Learning," in *Proc. Int. Conf. Robotics and Automation (ICRA)*, 2019, pp. 8543–8549.

- Typical visuomotor robot policies produce motor commands, but **have no notion about the quality of the selected motions under certain conditions**
  - Resulting policies may be unsafe as a result
- Representing uncertainty over the output requires **modelling the output as a distribution** rather than as a maximum likelihood point estimate — Bayesian inference is required
- There are various methods of encoding uncertainty in neural networks, most of them under the umbrella of **variational Bayes methods**

Representing Uncertainty in Policies



K. Lee, K. Saigol and E. A. Theodorou, "Early Failure Detection of Deep End-to-End Control Policy by Reinforcement Learning," in *Proc. Int. Conf. Robotics and Automation (ICRA)*, 2019, pp. 8543–8549.

- Typical visuomotor robot policies produce motor commands, but **have no notion about the quality of the selected motions under certain conditions**
  - Resulting policies may be unsafe as a result
- Representing uncertainty over the output requires **modelling the output as a distribution** rather than as a maximum likelihood point estimate — Bayesian inference is required
- There are various methods of encoding uncertainty in neural networks, most of them under the umbrella of **variational Bayes methods**
- Uncertainty information about the output can **enable safety behaviours to be triggered** — that can also include asking for human help