Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Institute for AI and Autonomous Systems

Learning-Based Object Grasping

An Overview

Dr. Alex Mitrevski
Master of Autonomous Systems

Structure

REVIEW article

Front. Robot. AI, 04 April 2023
Sec. Robot Learning and Evolution
Volume 10 - 2023 | https://doi.org/10.3389/frobt.2023.1038658

Learning-based robotic grasping: A review

Zhen Xie[1]* Xinquan Liang[2] Canale Roberto[1]

IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*

The International Journal of Robotics Research

Robot Grasp Synthesis Algorithms: A Survey

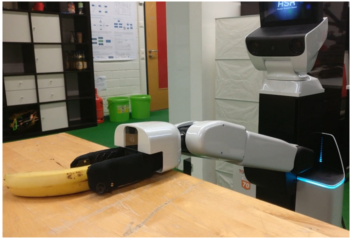
K.B. Shimoga View all authors and affiliations

Volume 15, Issue 3 | https://doi.org/10.1177/027836499601500302

- Object grasping primer
- Learning-based grasping
- A closer look at concrete learning-based grasping frameworks

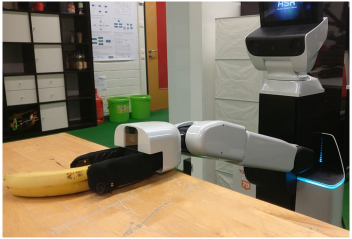
Object Grasping Primer

What is Object Grasping?



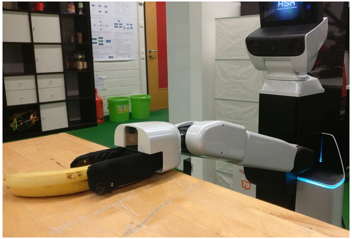
- Informally, object grasping is **the problem of picking up an object with a robotic hand**

What is Object Grasping?



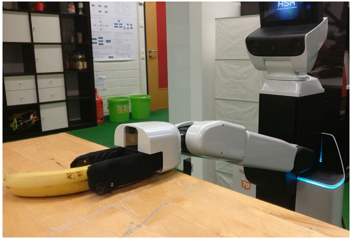
- Informally, object grasping is **the problem of picking up an object with a robotic hand**
- The problem can be defined as that of **finding an end-effector pose** $p^{G^*} = (t^{G^*}, R^{G^*})$ **that ensures a stable object grasp** (the grasped object will remain in the gripper)

What is Object Grasping?



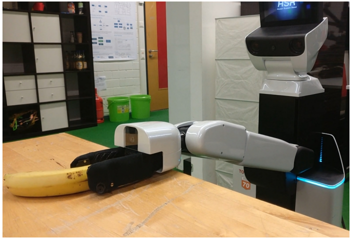
- Informally, object grasping is **the problem of picking up an object with a robotic hand**
- The problem can be defined as that of **finding an end-effector pose** $\boldsymbol{p}^{G^*} = (\boldsymbol{t}^{G^*}, R^{G^*})$ **that ensures a stable object grasp** (the grasped object will remain in the gripper)
  - Here, $\boldsymbol{t}^{G^*}$ is the grasp center point and R^{G^*} is the gripper's (wrist) orientation

What is Object Grasping?



- Informally, object grasping is **the problem of picking up an object with a robotic hand**
- The problem can be defined as that of **finding an end-effector pose $p^{G^*} = (t^{G^*}, R^{G^*})$ that ensures a stable object grasp** (the grasped object will remain in the gripper)
  - Here, t^{G^*} is the grasp center point and R^{G^*} is the gripper's (wrist) orientation
  - Often, a grasp is additionally parameterised by an approach vector $\boldsymbol{a}$

What is Object Grasping?



- Informally, object grasping is **the problem of picking up an object with a robotic hand**
- The problem can be defined as that of **finding an end-effector pose $\boldsymbol{p}^{G^*} = (\boldsymbol{t}^{G^*}, R^{G^*})$ that ensures a stable object grasp** (the grasped object will remain in the gripper)
  - Here, $\boldsymbol{t}^{G^*}$ is the grasp center point and R^{G^*} is the gripper's (wrist) orientation
  - Often, a grasp is additionally parameterised by an approach vector $\boldsymbol{a}$
- More generally, the problem is that of **determining positions and applied forces for the individual gripper fingers**; let us denote such a grasp candidate by C^*

Grasp Synthesis

- Grasp synthesis is **an optimisation process that generates grasp candidates which satisfy a set of desired grasp quality metrics**

Grasp Synthesis

- Grasp synthesis is **an optimisation process that generates grasp candidates which satisfy a set of desired grasp quality metrics**
- Formally, a grasp synthesis procedure **takes on an object model O and a gripper description Q** (consisting of finger joint positions and forces/torques) and **generates a grasp candidate C^* that optimises the quality metrics**

Grasp Synthesis

- Grasp synthesis is **an optimisation process that generates grasp candidates which satisfy a set of desired grasp quality metrics**
- Formally, a grasp synthesis procedure **takes on an object model O and a gripper description Q** (consisting of finger joint positions and forces/torques) and **generates a grasp candidate C^* that optimises the quality metrics**
- Grasp synthesis procedures are **often sampling-based**, namely n grasping candidates $C_i, 1 \leq i \leq n$ are generated and scored based on the desired quality metrics; the candidate C^* with the highest score is then selected for grasping

The International Journal of Robotics Research
Robot Grasp Synthesis Algorithms: A Survey
K.B. Shimoga View all authors and affiliations
Volume 15, Issue 3 | https://doi.org/10.1177/027836499601500302

Grasp Properties

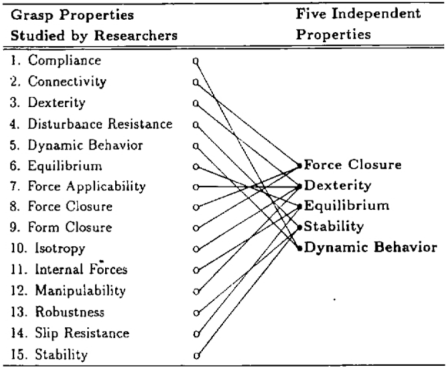
Grasp quality properties

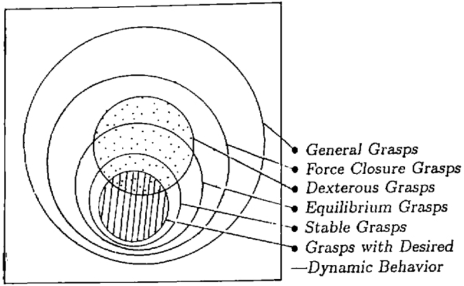
- **Dexterity**: The finger configuration should satisfy one or more dexterity measures, such as avoiding singularities
- **Equilibrium**: The forces and moments acting on the grasped object must add up to 0
- **Stability**: The grasped object must return to equilibrium once any external disturbances are removed
- **Dynamic behaviour**: The behaviour of the grasped object as a result of applied fingertip forces should follow a desired dynamic profile

The International Journal of Robotics Research
Robot Grasp Synthesis Algorithms: A Survey
K.B. Shimoga View all authors and affiliations
Volume 15, Issue 3 | https://doi.org/10.1177/027836499601500302

Grasp Quality Metrics

- There are a variety of **quality metrics** that can be used to evaluate grasps
- Such metrics are typically used in the **evaluation of grasp candidates proposed by grasp synthesis procedures**





Factors Affecting Grasp Synthesis

IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*

- In general, there are various factors that influence how grasp synthesis is performed

Factors Affecting Grasp Synthesis


IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*

- In general, there are various factors that influence how grasp synthesis is performed
- **Knowledge of an object (or lack thereof)** is one such factor — if an object is known, that knowledge can be exploited in the synthesis


IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*

Factors Affecting Grasp Synthesis

- In general, there are various factors that influence how grasp synthesis is performed
- **Knowledge of an object (or lack thereof)** is one such factor — if an object is known, that knowledge can be exploited in the synthesis
- Synthesis also depends on **the input modality used to identify objects** and **the types of object features** that are utilised by the candidate generation procedure

Factors Affecting Grasp Synthesis

- In general, there are various factors that influence how grasp synthesis is performed
- **Knowledge of an object (or lack thereof)** is one such factor — if an object is known, that knowledge can be exploited in the synthesis
- Synthesis also depends on **the input modality used to identify objects** and **the types of object features** that are utilised by the candidate generation procedure
- **The type of robotic hand for which grasps are generated** also influences how grasp candidates are generated

Factors Affecting Grasp Synthesis


IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*

- In general, there are various factors that influence how grasp synthesis is performed
- **Knowledge of an object (or lack thereof)** is one such factor — if an object is known, that knowledge can be exploited in the synthesis
- Synthesis also depends on **the input modality used to identify objects** and **the types of object features** that are utilised by the candidate generation procedure
- **The type of robotic hand for which grasps are generated** also influences how grasp candidates are generated
- **Task information** can also be useful to incorporate — this can constrain the set of valid grasp candidates

Challenges With Analytical Grasp Synthesis

Reliance on object models

Most analytical methods rely on given (geometric and physical) object models; this makes it **difficult to use them for unknown objects**

Challenges With Analytical Grasp Synthesis

Reliance on object models

Most analytical methods rely on given (geometric and physical) object models; this makes it **difficult to use them for unknown objects**

Reliance on simulations

Grasp quality metrics are often evaluated in simulation, but **simulated metrics do not necessarily translate well to the real world**

Challenges With Analytical Grasp Synthesis

Reliance on object models

Most analytical methods rely on given (geometric and physical) object models; this makes it **difficult to use them for unknown objects**

Reliance on simulations

Grasp quality metrics are often evaluated in simulation, but **simulated metrics do not necessarily translate well to the real world**

Slow synthesis

The evaluation of grasp quality metrics can be **computationally expensive**, which contributes to a slow synthesis process

Challenges With Analytical Grasp Synthesis

Reliance on object models

Most analytical methods rely on given (geometric and physical) object models; this makes it **difficult to use them for unknown objects**

Reliance on simulations

Grasp quality metrics are often evaluated in simulation, but **simulated metrics do not necessarily translate well to the real world**

Slow synthesis

The evaluation of grasp quality metrics can be **computationally expensive**, which contributes to a slow synthesis process

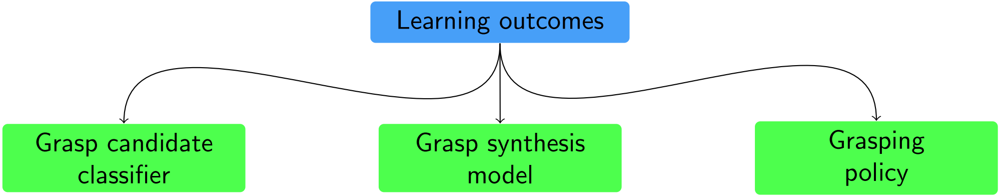
Inability to use prior experiences

Analytical approaches are **unable to use prior experiences to guide the synthesis process** — every grasping instance is treated independently

Learning-Based Grasping

Learning for Object Grasping

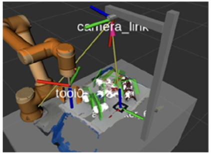
- The objective of learning for grasping is to **replace — partially or completely — analytical grasping methods**
- There are various ways in which this can be performed, depending on what the desired learning outcome is

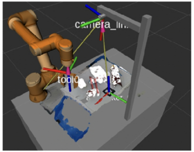


Grasp Candidate Classification

Front. Robot. AI, 04 April 2023
Sec. Robot Learning and Evolution
Volume 10 - 2023 | https://doi.org/10.3389/frobt.2023.1038658
Learning-based robotic grasping: A review
Zhen Xie[1]* Xinquan Liang[2] Canale Roberto[1]

- One way in which learning can be applied in the context of grasping is to **train a model that can be used to score grasp candidates** based on their quality

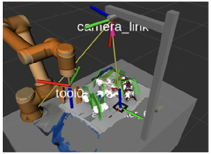


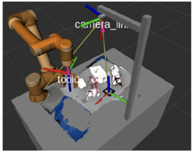


Front. Robot. AI, 04 April 2023
Sec. Robot Learning and Evolution
Volume 10 - 2023 | https://doi.org/10.3389/frobt.2023.1038658
Learning-based robotic grasping: A review
Zhen Xie[1]* Xinquan Liang[2] Canale Roberto[1]

Grasp Candidate Classification

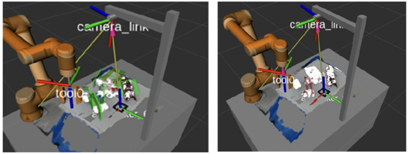
- One way in which learning can be applied in the context of grasping is to **train a model that can be used to score grasp candidates** based on their quality
- In this case, **a labelled dataset is needed** that contains ground-truth metrics as labels





Grasp Candidate Classification

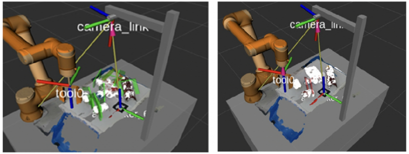
Front. Robot. AI, 04 April 2023
Sec. Robot Learning and Evolution
Volume 10 - 2023 | https://doi.org/10.3389/frobt.2023.1038658
Learning-based robotic grasping: A review
Zhen Xie[1]* Xinquan Liang[2] Canale Roberto[1]



- One way in which learning can be applied in the context of grasping is to **train a model that can be used to score grasp candidates** based on their quality
- In this case, **a labelled dataset is needed** that contains ground-truth metrics as labels
- **During online application, a set of grasp candidates needs to be generated**, all of which can be scored by the learned model

Grasp Candidate Classification

Front. Robot. AI, 04 April 2023
Sec. Robot Learning and Evolution
Volume 10 - 2023 | https://doi.org/10.3389/frobt.2023.1038658
Learning-based robotic grasping: A review
Zhen Xie[1]* Xinquan Liang[2] Canale Roberto[1]



- One way in which learning can be applied in the context of grasping is to **train a model that can be used to score grasp candidates** based on their quality
- In this case, **a labelled dataset is needed** that contains ground-truth metrics as labels
- **During online application, a set of grasp candidates needs to be generated**, all of which can be scored by the learned model
- **This strategy does not fully replace the analytical grasp pipeline** — analytical methods can still be used during grasp generation — but performs the grasp evaluation with a learned model
  - Analytical methods are sometimes used for generating the data labels

A Multitude of Learning-Based Grasp Evaluation Methods

- Grasp quality evaluation can be performed based on **various input modalities** and using **different evaluation models**
- An overview of some recent methods is given in the table on the right

Front. Robot. AI, 04 April 2023
Sec. Robot Learning and Evolution
Volume 10 - 2023 | https://doi.org/10.3389/frobt.2023.1038658

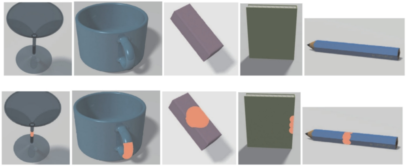
Learning-based robotic grasping: A review

Zhen Xie[1]* Xinquan Liang[2] Canale Roberto[1]

Grasping pose evaluation	Metric type	Input	Grasping type	Technical specification	Learning/ type
Yu et al. (2021a)	Support vector machine (SVM)	Superquadric shape parameters and grasping parameters	Hands, grippers, and suction	Difficult to represent complex shapes in superquadric parameters and vector sets as alternatives	Regression and supervised learning
Yen-Chen et al. (2021) and Irshad et al. (2022)	Analysis-by-synthesis optimization (AbS)	3D textured reconstruction	Hands, grippers, and suction	Complex multi-object scenarios and learned latent space	Regression
Katyara et al. (2021)	Kernel density estimation	Object-relative grasp poses	Hands, grippers, and suction	Continuous probability density functions, non-parametric	Non-parametric regression
MahlerLiang et al. (2017) and Zhang et al. (2021)	Robust grasp planning (RGP)	Prior grasps and 3D object models	Hands, grippers, and suction	Correlated bandit techniques and cloud-based object models	Multi-armed bandit and CNN
Liu et al. (2022a)	Force closure	Object and gripper pose, contact, and friction	Hands, grippers, and soft	Reduces the complexity and universal, force spiral space, and binary	N/A
de Souza et al. (2021)	Grasp wrench space analysis (GWS)	Contact location, contact normal, and frictional coefficients (Weisz and Allen, 2012)	Hands, grippers, and soft	Epsilon quality and magnitude of the minimum norm wrench	Deep learning
Wang et al. (2022a)	Variational autoencoder (VAE)	Primitive grasp set with the generated grasp set and gripper configuration	Hands, grippers, and suction	Compressed representation, Kullback–Leibler (KL) divergence, and latent space sampling	CNN and machine learning
HuangNagaraj et al. (2021)	Random forest	Grasp features	Hands, grippers, and suction	Quantified as its Gini impurity-based importance can be used for deformable grasping	Supervised machine learning
Jiang et al. (2022)	Grasp quality convolutional neural network (GQ-CNN)	Point clouds, grasps, and analytic grasp metrics MahlerLiang et al. (2017)	Hands, grippers, and suction	Grasp features represented as the angle, planar position, and depth of a gripper relative to an RGB-D camera	CNN
Wang et al. (2022b)	Deep geometry-aware grasping network (DGGN) Yan et al. (2018)	Point cloud, shape, location, and orientation	Parallel jaw grippers	Shape generation network and grasping outcome prediction network	Deep Learning and 3D CNN
Mi et al. (2021) and Liang and Boularias (2022)	Dynamic graph CNN	Segmented depth and color image	Hands, grippers, and suction	Generalizes to new objects with different geometries and textures	CNN
Yu et al. (2021b)	Cascaded architecture of random forests Asif et al. (2017a) and Asif et al. (2017b)	RGB-D point clouds	Hands, grippers, and suction	Object-class and grasp-pose probabilities are computed, separated, and fused for unknown objects	CNN
Ayoobi et al. (2022)	Supervised bag-of-visual-words	Scene data	Hands, grippers, and suction	Uses local feature descriptors to match database Ergene and Durdu (2017)	Supervised learning
Miften et al. (2021)	AdaBoost	Object shape, grasp information, tactile information, and gripper configuration	Hands, grippers, and suction	Probabilistic learning framework, capable of inferring based on tactile measurement	Ensemble learning
Jiang et al. (2021)	Grasping pose detection (GDP)	Point cloud and gripper configuration	Hands, grippers, and suction	Directly on the point cloud w/o estimating grasping pose, can be used in the clustered environment	Deep learning

Learning-Based Grasp Synthesis

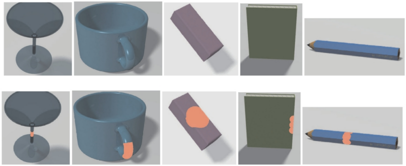
- Another way to apply learning for grasping is to **learn a model that can be used to identify suitable grasping candidates**



A. Saxena, J. Driemeyer, and A. Y. Ng. “Robotic grasping of novel objects using vision,” *International Journal of Robotics Research*, vol. 27, no. 2, pp. 157–173, 2008.

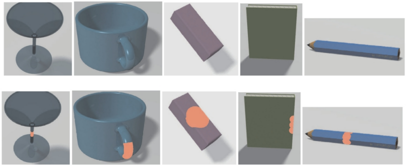
Learning-Based Grasp Synthesis

- Another way to apply learning for grasping is to **learn a model that can be used to identify suitable grasping candidates**
- In this case, learning is typically done by **mapping visual (or multimodal) object features to a grasp quality estimate**



A. Saxena, J. Driemeyer, and A. Y. Ng. "Robotic grasping of novel objects using vision," *International Journal of Robotics Research*, vol. 27, no. 2, pp. 157–173, 2008.

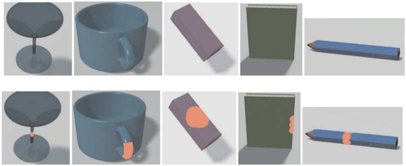
Learning-Based Grasp Synthesis



A. Saxena, J. Driemeyer, and A. Y. Ng. "Robotic grasping of novel objects using vision," *International Journal of Robotics Research*, vol. 27, no. 2, pp. 157–173, 2008.

- Another way to apply learning for grasping is to **learn a model that can be used to identify suitable grasping candidates**
- In this case, learning is typically done by **mapping visual (or multimodal) object features to a grasp quality estimate**
- The online application of such a model involves **feature estimation** and then **mapping those features through the learned model so that grasp candidates are found**

Learning-Based Grasp Synthesis



A. Saxena, J. Driemeyer, and A. Y. Ng. "Robotic grasping of novel objects using vision," *International Journal of Robotics Research*, vol. 27, no. 2, pp. 157–173, 2008.

- Another way to apply learning for grasping is to **learn a model that can be used to identify suitable grasping candidates**
- In this case, learning is typically done by **mapping visual (or multimodal) object features to a grasp quality estimate**
- The online application of such a model involves **feature estimation** and then **mapping those features through the learned model so that grasp candidates are found**
- The use of analytical grasp quality metrics is less common in this case — labels are typically simpler and denote whether a point represents a valid grasp or not

Grasp Policies

- The previous two techniques generate grasp candidates or evaluate candidates, but **grasping can also be performed using a learned policy, without explicit grasp hypotheses**

Grasp Policies

- The previous two techniques generate grasp candidates or evaluate candidates, but **grasping can also be performed using a learned policy, without explicit grasp hypotheses**
- In this case, **a visuomotor policy for object grasping is learned**

Grasp Policies

- The previous two techniques generate grasp candidates or evaluate candidates, but **grasping can also be performed using a learned policy, without explicit grasp hypotheses**
- In this case, **a visuomotor policy for object grasping is learned**
- During online application, **the policy can be directly applied without extracting explicit grasp candidates**

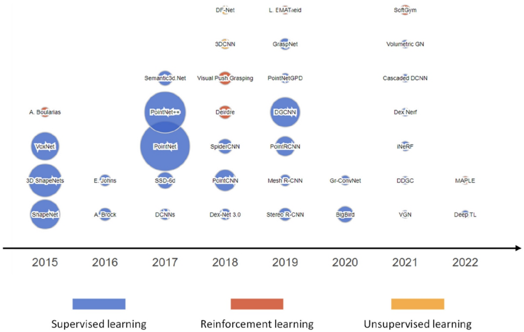
Grasp Policies

- The previous two techniques generate grasp candidates or evaluate candidates, but **grasping can also be performed using a learned policy, without explicit grasp hypotheses**
- In this case, **a visuomotor policy for object grasping is learned**
- During online application, **the policy can be directly applied without extracting explicit grasp candidates**
- Such a policy can be trained with a **sparse reward** (reward only on successful grasping) or with a **shaped reward** (where the shaping could potentially use analytical grasping metrics)

Front. Robot. AI, 04 April 2023
Sec. Robot Learning and Evolution
Volume 10 - 2023 | https://doi.org/10.3389/frobt.2023.1038658
Learning-based robotic grasping: A review
Zhen Xie[1]* Xinquan Liang[2] Canale Roberto[1]

Techniques Used in Learning-Based Grasping

- The previously mentioned learning problems can be solved using any of the usual learning paradigms
- **Supervised learning appears to be the predominant learning paradigm in the literature**
- A variety of supervised learning techniques are **particularly used for feature extraction**, for instance PointNet++ and DGCNN for extracting features from point cloud data
- Existing databases of 3D object models, such as ShapeNet, as well as point cloud datasets, such as Semantic3D, have been applied in this context



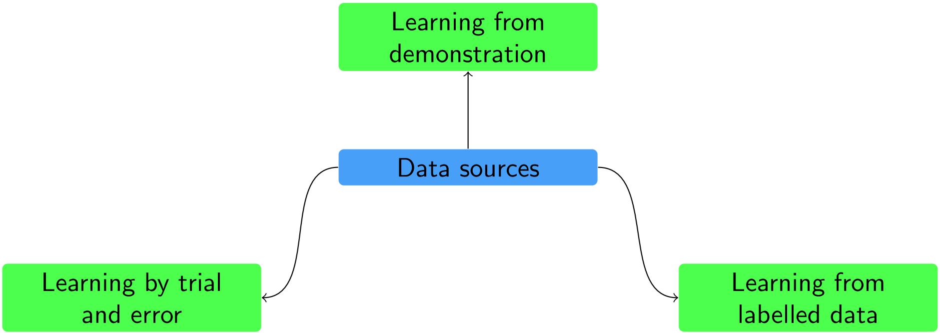
Hochschule Bonn-Rhein-Sieg University of Applied Sciences
b-it Bonn-Aachen International Center for Information Technology
Institute for AI and Autonomous Systems

Learning Data Sources


IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*



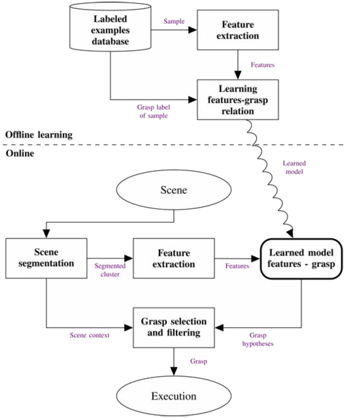
Learning from Labelled Data


IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*

- Supervised learning is the most common strategy applied in the context of grasping
- A variety of models can be learned here, for instance:
  - **A grasp candidate evaluation model**
  - **A database of grasp samples**
  - **A grasp sampling model**
- Since obtaining enough labelled data for learning can be challenging, **synthethic data** are sometimes used for learning



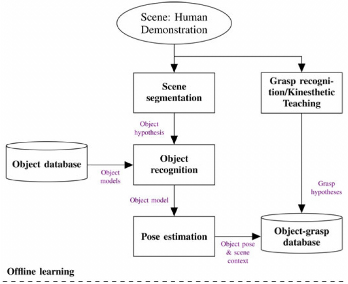
Learning from Demonstrations

IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*

- Demonstrations represent another potential data source for grasp learning
- The objective in this case is to obtain a **small set of successful grasps** that can be used directly for execution or for further grasp model learning
- Demonstrations can be used **as an independent data source** or can **supplement data used for supervised learning**



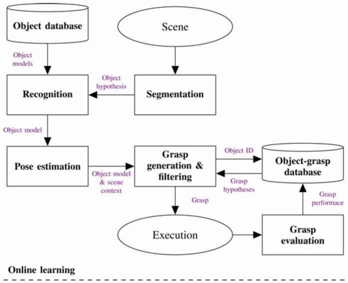
Learning by Trial and Error

IEEE TRANSACTIONS ON ROBOTICS, VOL. 30, NO. 2, APRIL 2014 289

Data-Driven Grasp Synthesis—A Survey

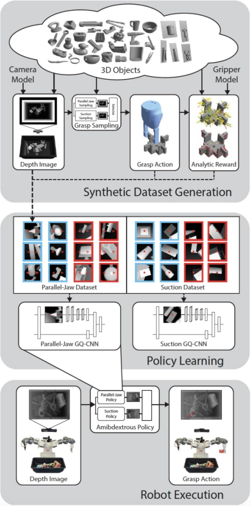
Jeannette Bohg, *Member, IEEE*, Antonio Morales, *Member, IEEE*, Tamim Asfour, *Member, IEEE*, and Danica Kragic, *Senior Member, IEEE*

- If neither labelled data nor human demonstrations are available, **a robot can collect its own data for learning grasping models**
- The nature of the collected data would differ depending on whether a grasp synthesis model or a grasp policy is learned
- If the collected data includes a mapping between object identities and attempted grasps, **efficient object-specific grasping models can be acquired**

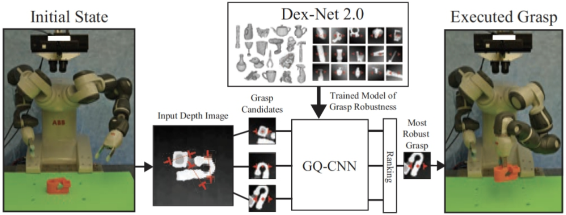


A Closer Look at Concrete Learning-Based Grasping Frameworks

Dexterity Network (Dex-Net)



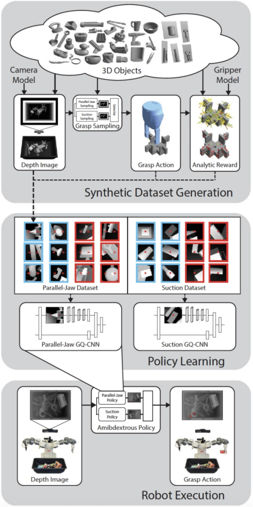
J. Mahler et al., "Learning ambidextrous robot grasping policies," *Science Robotics*, vol. 4, no. 26, p. eaau4984, 2019.



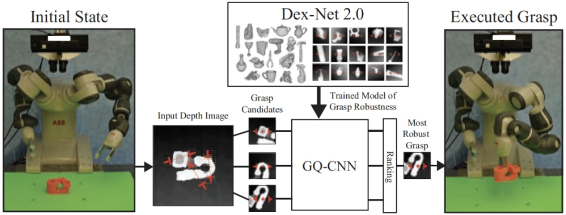
J. Mahler et al., "Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics," in *Proc. Robotics: Science and Systems (RSS)*, 2017.

- Dex-Nex is a **convolutional neural network-based grasp quality evaluation model**

Dexterity Network (Dex-Net)



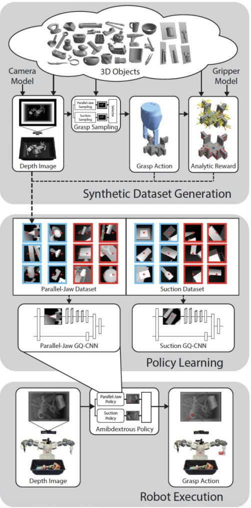
J. Mahler et al., "Learning ambidextrous robot grasping policies," *Science Robotics*, vol. 4, no. 26, p. eaau4984, 2019.



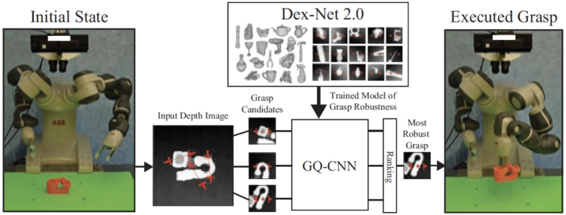
J. Mahler et al., "Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics," in *Proc. Robotics: Science and Systems (RSS)*, 2017.

- Dex-Nex is a **convolutional neural network-based grasp quality evaluation model**
- The earlier Dex-Net versions are defined for parallel-jaw grippers; the latest version can also be used with suction grippers

Dexterity Network (Dex-Net)



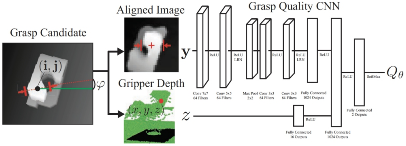
J. Mahler et al., "Learning ambidextrous robot grasping policies," *Science Robotics*, vol. 4, no. 26, p. eaau4984, 2019.

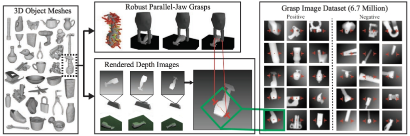


J. Mahler et al., "Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics," in *Proc. Robotics: Science and Systems (RSS)*, 2017.

- Dex-Nex is a **convolutional neural network-based grasp quality evaluation model**
- The earlier Dex-Net versions are defined for parallel-jaw grippers; the latest version can also be used with suction grippers
- Dex-Net's parallel-jaw grasps are parameterised by:
  - **pixel coordinates** (determine a grasp position considering a top-down object view)
  - **gripper depth** (grasping height)
  - **gripper orientation**

Grasp Quality Convolutional Neural Network (GQ-CNN)

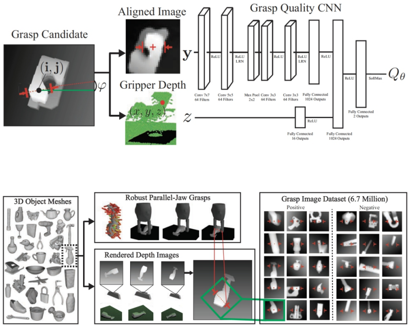




J. Mahler et al., "Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics," in *Proc. Robotics: Science and Systems (RSS)*, 2017.

- The main element behind Dex-Net is a **Grasp Quality Convolutional Neural Network (GQ-CNN)**

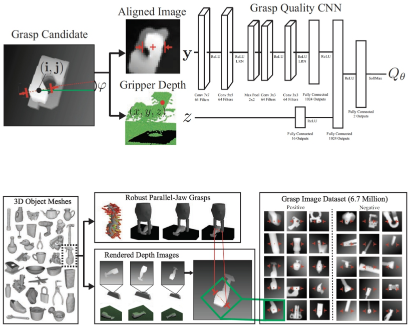
Grasp Quality Convolutional Neural Network (GQ-CNN)



J. Mahler et al., "Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics," in *Proc. Robotics: Science and Systems (RSS)*, 2017.

- The main element behind Dex-Net is a **Grasp Quality Convolutional Neural Network (GQ-CNN)**
- The network takes as input **a grasp candidate represented as an aligned depth image** and **the gripper depth**, and outputs **an estimate of the grasp success probability**

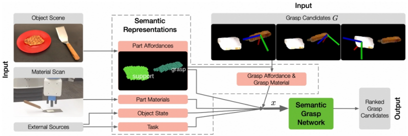
Grasp Quality Convolutional Neural Network (GQ-CNN)

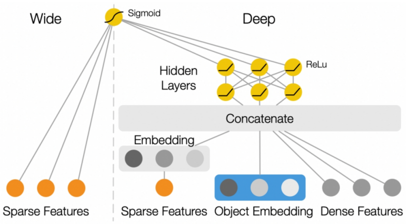


J. Mahler et al., "Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics," in *Proc. Robotics: Science and Systems (RSS)*, 2017.

- The main element behind Dex-Net is a **Grasp Quality Convolutional Neural Network (GQ-CNN)**
- The network takes as input **a grasp candidate represented as an aligned depth image** and **the gripper depth**, and outputs **an estimate of the grasp success probability**
- An important aspect about GQ-CNN is that it is **trained using synthetically generated data**
  - The synthetic dataset is created using **3D object models**
  - Grasp candidates for training are generated by **sampling latent variables from a graphical model**
  - For grasp samples in the training data, **a success metric (probability of force closure) is evaluated** in order to generate a label for the candidate

Context-Aware Grasping

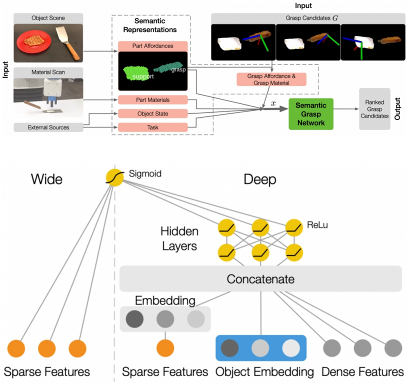




W. Liu, A. Daruna and S. Chernova, "CAGE: Context-Aware Grasping Engine," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020, pp. 2550-2556.

- The CAGE model is a **deep neural network that evaluates grasp candidates**, calculating a likelihood that the candidate would lead to a successful grasp

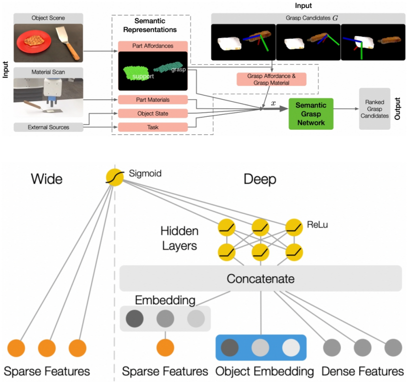
Context-Aware Grasping



W. Liu, A. Daruna and S. Chernova, "CAGE: Context-Aware Grasping Engine," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020, pp. 2550-2556.

- The CAGE model is a **deep neural network that evaluates grasp candidates**, calculating a likelihood that the candidate would lead to a successful grasp
- The model uses information about the **task context**, which combines multiple elements:
  - semantic task information (one-hot task and object state encodings)
  - affordance estimation for a grasp candidate point
  - point material information

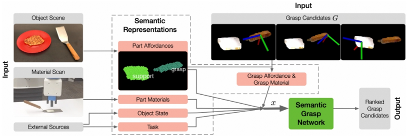
Context-Aware Grasping

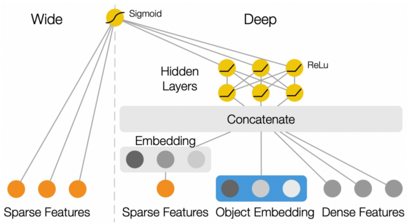


W. Liu, A. Daruna and S. Chernova, "CAGE: Context-Aware Grasping Engine," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020, pp. 2550-2556.

- The CAGE model is a **deep neural network that evaluates grasp candidates**, calculating a likelihood that the candidate would lead to a successful grasp
- The model uses information about the **task context**, which combines multiple elements:
  - semantic task information (one-hot task and object state encodings)
  - affordance estimation for a grasp candidate point
  - point material information
- The network is based on a **wide-and-deep architecture** in which:
  - the wide component processes the task context
  - the deep component combines the task context, an object embedding, and optional additional features

Context-Aware Grasping





W. Liu, A. Daruna and S. Chernova, "CAGE: Context-Aware Grasping Engine," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2020, pp. 2550-2556.

- The CAGE model is a **deep neural network that evaluates grasp candidates**, calculating a likelihood that the candidate would lead to a successful grasp
- The model uses information about the **task context**, which combines multiple elements:
  - semantic task information (one-hot task and object state encodings)
  - affordance estimation for a grasp candidate point
  - point material information
- The network is based on a **wide-and-deep architecture** in which:
  - the wide component processes the task context
  - the deep component combines the task context, an object embedding, and optional additional features
- Training is done with a **negative log-likelihood loss**