SAIR

Spatial AI & Robotics Lab

CSE 473/573-A

L27: POINT CLOUD REGISTRATION

Qiwei Du

Spatial AI & Robotics Lab

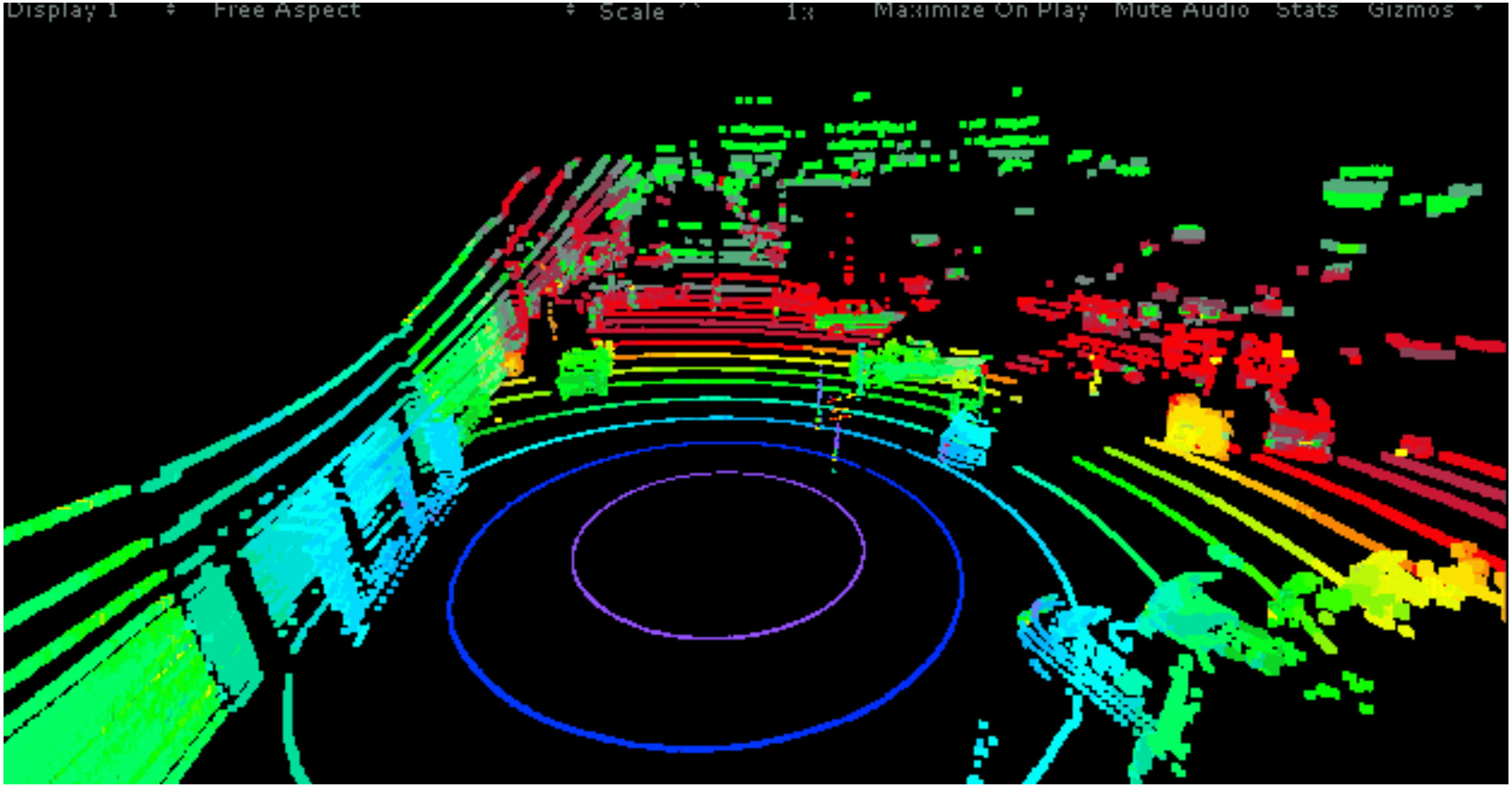
Department of Computer Science and Engineering

University at Buffalo The State University of New York

Outline

- Application of Point Cloud Registration (PCR)
- Definition of PCR
- Iterative Closest Point (ICP)
- Feature Matching Based Methods
- End-to-end Deep Learning Methods

Why Should We Care About PCR?



More Applications of PCR

- Robotics and Autonomous Navigation
  - SLAM
  - Collision Avoidance
  - Path Planning
- Augmented Reality (AR) and Virtual Reality (VR)
  - Scene Reconstruction
  - Object Placement
- Geospatial Mapping and Surveying
  - Aerial Mapping
  - Urban Planning
  - Change Detection

More Applications of PCR

- Medical Imaging
  - Surgical Planning
  - Disease Progression Monitoring
  - Prosthetic Design
- Cultural Heritage Preservation
  - Artifact Reconstruction
  - Site Documentation
  - Virtual Tours
- Manufacturing and Quality Control
  - 3D Inspection
  - Assembly Line Automation
  - Reverse Engineering

More Applications of PCR

- Construction and Building Information Modeling
  - Structural Analysis
  - Retrofitting and Renovation
  - Deformation Monitoring
- Environmental Monitoring
  - Forest Canopy Analysis
  - Coastal Erosion Tracking
  - Glacier Movement
- Entertainment and Media
  - 3D Modeling for Films and Games
  - Motion Capture
  - Scene Integration

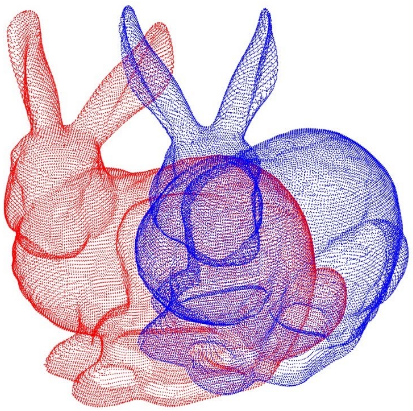
More Applications of PCR

- Disaster Response and Management
  - Damage Assessment
  - Search and Rescue
  - Flood Modeling
- ……

Definition of Point Cloud Registration

- Goal
  - Find the parameters of the transformation that best align corresponding data points

- Methods
  - Iterative Closest Point (ICP)
  - Feature Matching Based
    - FPFH, D3Feat, Predator, …
  - End-to-end
    - DCP, RegTR, …
  - ……



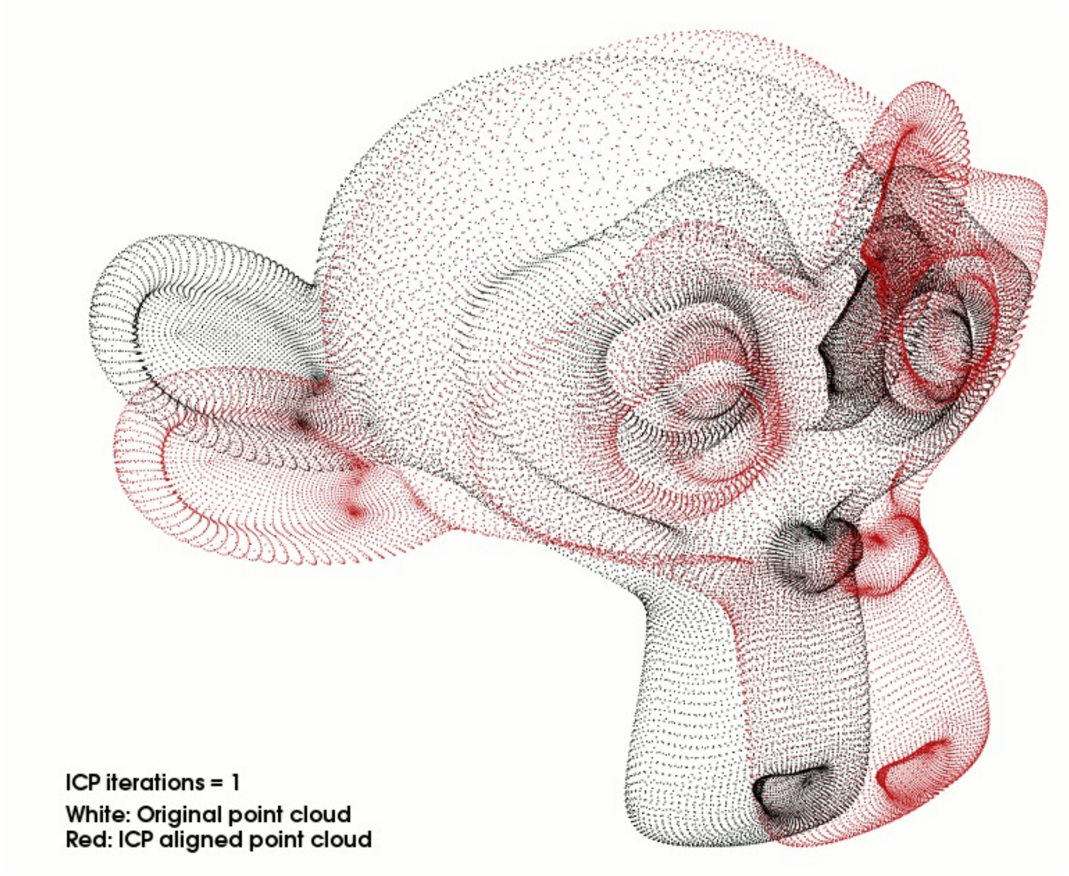
Iterative Closest Point (ICP)

- Transformation Parameters (rigid)
  - A translation vector
  - A rotation matrix

- Calculate with Known correspondences

- Calculate with Estimated correspondences

Find Relative Transformation



Basic Registration Problem

- Given two input point sets:

$$\text{source point cloud } \mathbf{X} \in \mathbb{R}^{N \times 3}$$

$$\text{target point cloud } \mathbf{Y} \in \mathbb{R}^{M \times 3}$$

  - with correspondences: C = { (i, j) }

- Find:

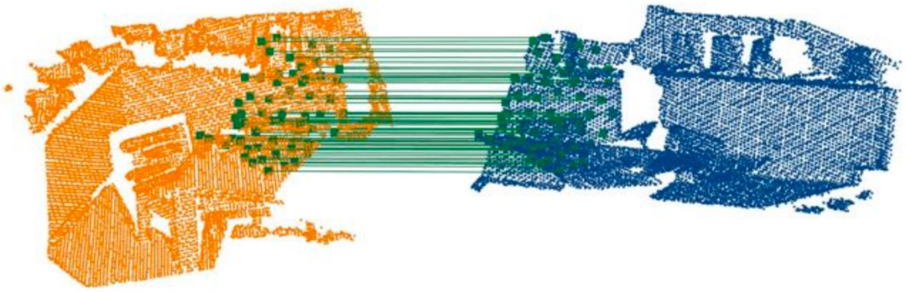
$$\text{a rigid transformation } \mathbf{T} = \{\mathbf{R} \in \mathrm{SO}(3), \mathbf{t} \in \mathbb{R}^3\}$$

  - that minimize:

$$\sum_{i}^{M'+N'} \|\mathbf{R}\hat{\mathbf{x}}_i + \mathbf{t} - \hat{\mathbf{y}}_i\|^2$$

Key Idea

- Given correct correspondences, find the rotation and translation in one go with the direct solution.



Only focus on points with correspondence

- Assume: n pairs in correspondence set C = { (i, j) }
- Reorder point cloud $X = \{x_n\}$, $Y = \{y_n\}$
- Find R, t, so that $\{R \cdot x_n + t\}$ is as close to $\{y_n\}$ as possible
  - Minimize the sum of square point-to-point distances

Steps of calculating (R, t) with GT pairs

1. **Compute Centroids**: Compute the centroids $\mathbf{x}_{\text{mean}}$ and $\mathbf{y}_{\text{mean}}$ of X and Y:

$$\mathbf{x}_{\text{mean}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i, \quad \mathbf{y}_{\text{mean}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{y}_i$$

2. **Center the Points**: Subtract the centroids to center the point clouds:

$$\mathbf{x}'_i = \mathbf{x}_i - \mathbf{x}_{\text{mean}}, \quad \mathbf{y}'_i = \mathbf{y}_i - \mathbf{y}_{\text{mean}}$$

3. **Compute the Covariance Matrix**: Construct the covariance matrix H using the centered points:

$$H = \sum_{i=1}^{N}\mathbf{x}'_i(\mathbf{y}'_i)^\top$$

4. **Perform Singular Value Decomposition (SVD)**: Decompose H using SVD:

$$H = U\Sigma V^\top$$

5. **Compute Rotation R**: The optimal rotation matrix is:

$$R = VU^\top$$

If $\det(R) < 0$, ensure a proper rotation by flipping the sign of the last column of V.

6. **Compute Translation t**: Finally, compute the translation vector:

$$t = \mathbf{y}_{\text{mean}} - R\mathbf{x}_{\text{mean}}$$

Complete steps of ICP

1. **Initialization**:
   - Begin with an initial guess for the transformation T, typically the identity transformation.
2. **Iterative Process**:
   - **Correspondence Estimation**: For each point $\mathbf{x}_i \in X$, find its closest point $\mathbf{y}_j \in Y$ using a distance metric, such as Euclidean distance.
   - **Transformation Estimation**: Using the matched pairs, compute R and t following the method described earlier.
   - **Apply Transformation**: Update X by applying T:

$$X \leftarrow \{R\mathbf{x}_i + t \mid \mathbf{x}_i \in X\}$$

   - **Check Convergence**: Measure the alignment error, and terminate if it falls below a threshold or if the change between iterations is negligible.
3. **Output**:
   - The final transformation T aligns X to Y.

Key Considerations and Challenges of ICP

- **Initialization**: A good initial guess for T can significantly improve convergence.
- **Outliers**: Outliers in the data can cause errors in correspondence estimation. Techniques like outlier rejection or weighted ICP help mitigate this issue.
- **Efficiency**: Computing nearest neighbors naively is computationally expensive. Efficient implementations use KD-trees or similar data structures for acceleration.

More categories of advanced PCR methods

- **Feature Matching-Based Methods**: These rely on finding discriminative features in the point clouds to establish correspondences.
- **End-to-End Deep Learning Methods**: These directly predict transformations without requiring explicit feature matching.

Feature Matching-Based Methods

- **Fast Point Feature Histograms (FPFH)**
  - Rusu, Radu Bogdan, Nico Blodow, and Michael Beetz. "Fast point feature histograms (FPFH) for 3D registration." In *2009 IEEE international conference on robotics and automation*, pp. 3212-3217. IEEE, 2009.

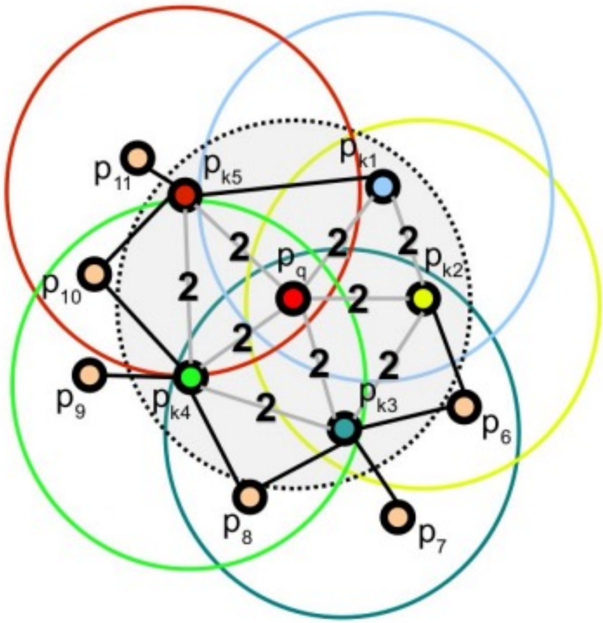


Fig. 5. The influence region diagram for a Fast Point Feature Histogram. Each query point (red) is connected only to its direct k-neighbors (enclosed by the gray circle). Each direct neighbor is connected to its own neighbors and the resulted histograms are weighted together with the histogram of the query point to form the FPFH. The connections marked with 2 will contribute to the FPFH twice.

SAIR
Spatial AI & Robotics Lab

Feature Matching-Based Methods

- **D3Feat**
  - Bai, Xuyang, Zixin Luo, Lei Zhou, Hongbo Fu, Long Quan, and Chiew-Lan Tai. "D3feat: Joint learning of dense detection and description of 3d local features." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 6359-6367. 2020.

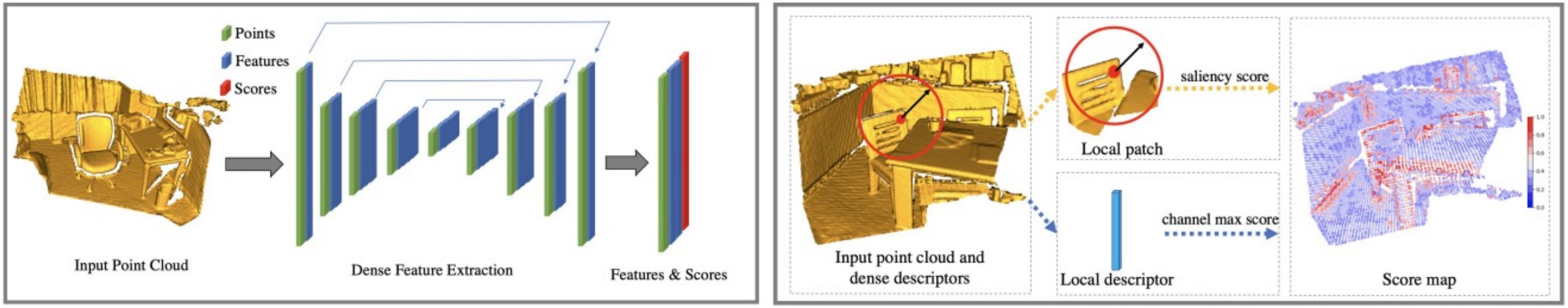


Figure 1: (Left) The network architecture of D3Feat. Each block indicates a ResNet block using KPConv to replace image convolution. All layers except the last one are followed by batch normalization and ReLU. (Right) Keypoint detection. After dense feature extraction, we calculate the keypoint detection scores by applying saliency score and channel max score. This figure is best viewed with color and zoom-in.

Feature Matching-Based Methods

- **Predator**
  - Huang, Shengyu, Zan Gojcic, Mikhail Usvyatsov, Andreas Wieser, and Konrad Schindler. "Predator: Registration of 3d point clouds with low overlap." In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pp. 4267-4276. 2021.

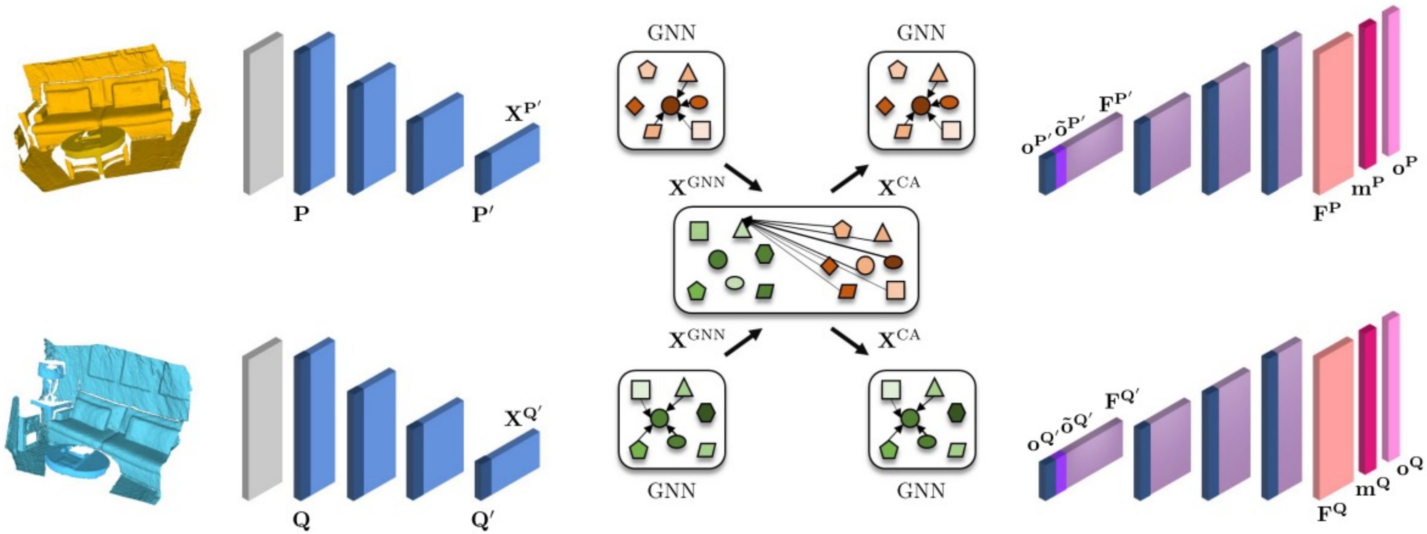


Figure 3: Network architecture of PREDATOR. Voxel-gridded point clouds $\mathbf{P}$ and $\mathbf{Q}$ are fed to the encoder, which extracts the superpoints $\mathbf{P}'$ and $\mathbf{Q}'$ and their latent features $\mathbf{X}^{\mathbf{P}'}$, $\mathbf{X}^{\mathbf{Q}'}$. The overlap-attention module updates the features with co-contextual information in a series of self- (GNN) and cross-attention (CA) blocks, and projects them to overlap $\mathbf{o}^{\mathbf{P}'}$, $\mathbf{o}^{\mathbf{Q}'}$ and cross-overlap $\tilde{\mathbf{o}}^{\mathbf{P}'}$, $\tilde{\mathbf{o}}^{\mathbf{Q}'}$ scores. Finally, the decoder transforms the conditioned features and overlap scores to per-point feature descriptors $\mathbf{F}^{\mathbf{P}}$, $\mathbf{F}^{\mathbf{Q}}$, overlap scores $\mathbf{o}^{\mathbf{P}}$, $\mathbf{o}^{\mathbf{Q}}$, and matchability scores $\mathbf{m}^{\mathbf{P}}$, $\mathbf{m}^{\mathbf{Q}}$.

Feature Matching-Based Methods

- **Other methods**
  - **Spin Images**: Another classical descriptor for capturing local geometry.
    - He, Yuqing, and Yuangang Mei. "An efficient registration algorithm based on spin image for LiDAR 3D point cloud models." *Neurocomputing* 151 (2015): 354-363.
  - **SHOT (Signature of Histograms of Orientations)**: Combines surface normal histograms with rotational invariance.
    - Salti, Samuele, Federico Tombari, and Luigi Di Stefano. "SHOT: Unique signatures of histograms for surface and texture description." *Computer Vision and Image Understanding* 125 (2014): 251-264.
  - **SuperPoint-like Approaches**: Adapt concepts from image processing to extract keypoints and features in point clouds.
    - DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Superpoint: Self-supervised interest point detection and description." In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 224-236. 2018.

End-to-End Deep Learning Methods

- **Deep Closest Point (DCP)**
  - Wang, Yue, and Justin M. Solomon. "Deep closest point: Learning representations for point cloud registration." In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3523-3532. 2019.

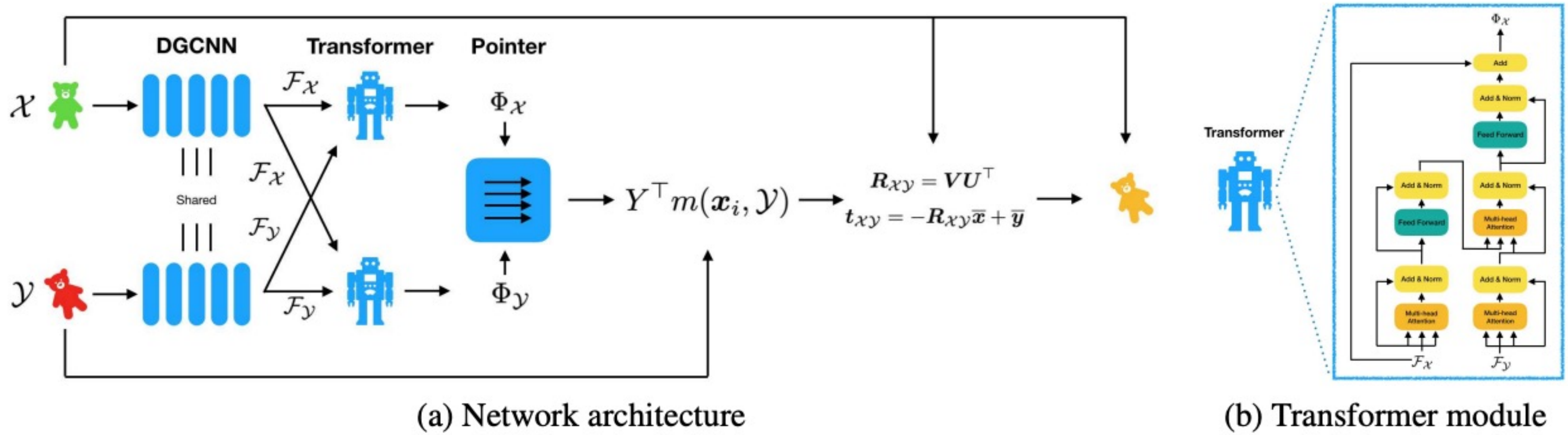


Figure 2. Network architecture for DCP, including the Transformer module for DCP-v2.

End-to-End Deep Learning Methods

- **RegTR**
  - Yew, Zi Jian, and Gim Hee Lee. "Regtr: End-to-end point cloud correspondences with transformers." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 6677-6686. 2022.

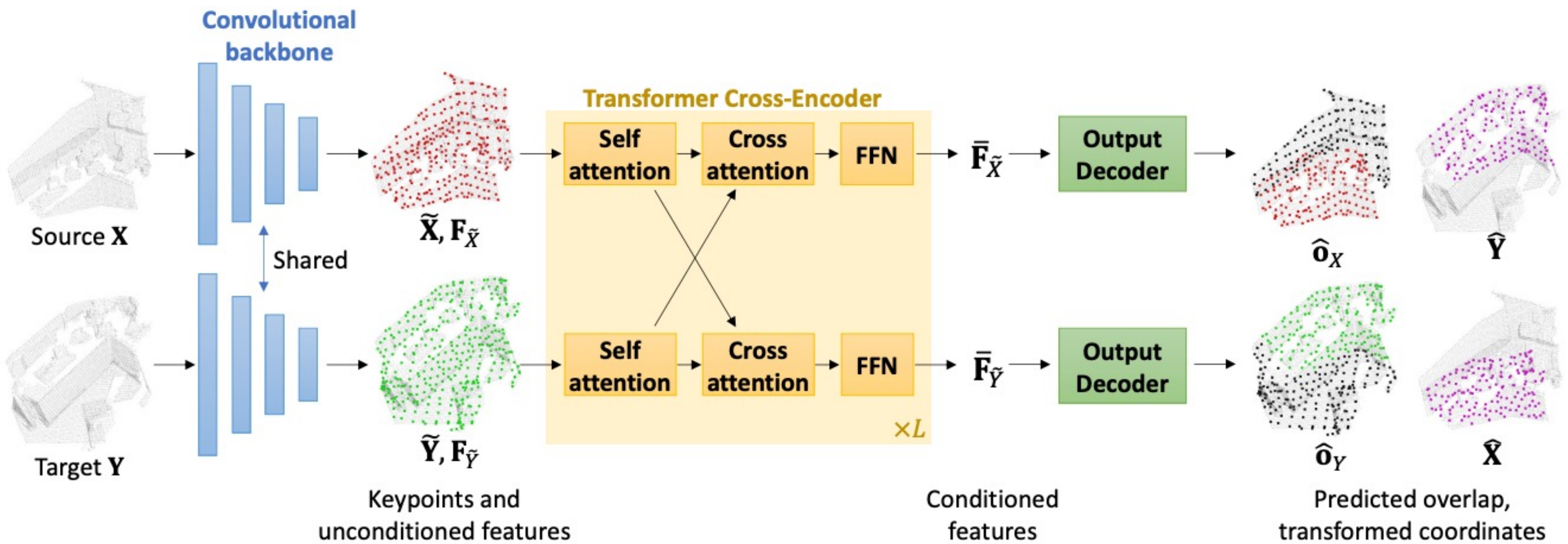


Figure 2. REGTR uses the KPConv convolutional backbone to extract a set of features for a sparse set of points. The features are then passed into several transformer cross-encoder layers. Lastly, the output decoder predicts the overlap score and the corresponding transformed coordinates of the sparse keypoints, which can be used for direct estimation of the pose. Best viewed in color.

SAIR
Spatial AI & Robotics Lab

End-to-End Deep Learning Methods

- **Other methods**
  - **PRNet**: Uses PointNet for feature extraction and directly predicts transformations.
    - Wang, Yue, and Justin M. Solomon. "Prnet: Self-supervised learning for partial-to-partial registration." *Advances in neural information processing systems* 32 (2019).
  - **PointDSC**: Integrates spatial consistency checks into the learning pipeline, improving robustness to noise.
    - Bai, Xuyang, Zixin Luo, Lei Zhou, Hongkai Chen, Lei Li, Zeyu Hu, Hongbo Fu, and Chiew-Lan Tai. "Pointdsc: Robust point cloud registration using deep spatial consistency." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 15859-15869. 2021.
  - **RPM-Net**: Combines feature learning with a probabilistic model for robust matching.
    - Yew, Zi Jian, and Gim Hee Lee. "Rpm-net: Robust point matching using learned features." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11824-11833. 2020.

Good luck with your final exams!