Read Group Meeting

Yuchen Li Vision-Cair Group

April 7th 2021

Few-shot Semantic Segmentation with Prototype Learning

Dong, Nanqing, and Eric P. Xing. "Few-Shot Semantic Segmentation with Prototype Learning." BMVC. Vol. 3. No. 4. 2018.

Introduction

Introduction

1. How to transplant an effective model for classification tasks to semantic segmentation tasks?

   The transplantation algorithm chosen by the author here is the prototype network Prototypical network. The entire network architecture consists of two parts: a prototype learner that learns from pictures and annotations to prototype and an annotation network that is used to annotate new pictures.

2. How to achieve one-to-many semantic segmentation?

   In order to solve the second problem, the author uses a data enhancement technique called **permutation training** to complete one-to-many semantic segmentation.

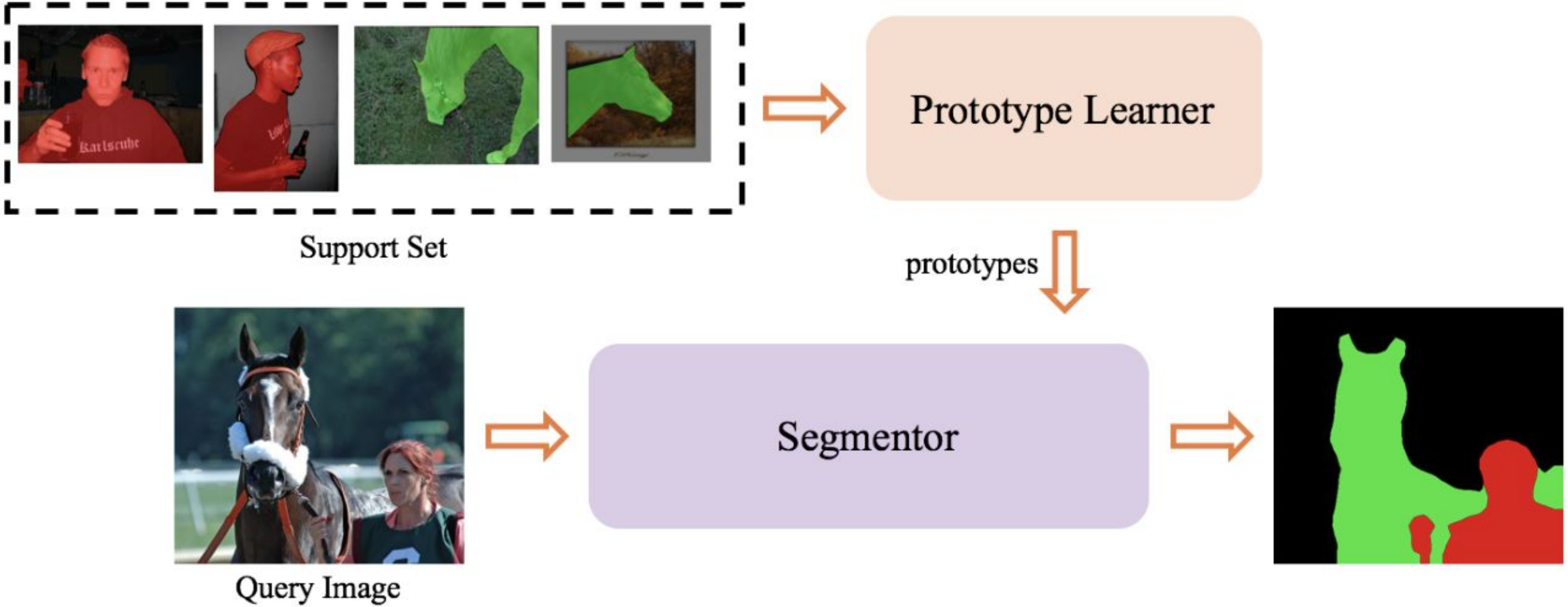


Figure 1: Illustration of a 2-way 2-shot semantic segmentation task. The prototype learner learns the prototypes from the support set and outputs the prototypes to the segmentor. The segmentor takes the query image and the prototypes to predict the segmentation mask. The 1-shot task can be easily extended to k-shot tasks by having k examples for each class in the support set.

Relative work

Semantic Segmentation

Semantic segmentation is the task of associating each pixel of an image with a semantic class label. Semantic segmentation can also be seen as a combination of the **semantic feature extraction task** and the **pixel-wise classification** task.

Fully Convolutional Networks (FCNs) have been the backbone architectures in many semantic segmentation tasks. However, similar to other data driven deep learning methods, FCN-based semantic segmentation models usually require large amounts of annotated data. We use **FCNs** as the backbone models to test the few-shot performance of the proposed framework.

Semantic Segmentation Example

Foreground background Segmentation

Foreground-background (FG-BG) segmentation is the task to find the foreground pixels with features different from the background pixels. FG-BG has played an important role in the pre-processing step for object detection, face detection and motion detection.

In a few-shot learning setting, increased model complexity and dependency on the training data may lead to overfitting. They fine-tune FGBG model with pre-trained weights as a strong baseline model for 1-way learning tasks.

Few Shot Classification

- Shaban et al. first propose a meta-learning method which uses a meta-learner to learn the small subset of parameters for the learner.

- . Shaban et al. also adapt a Siamese Network to few-shot semantic segmentation with a learned L1 distance for pixel-wise cross similarity, but the performance is worse than the meta learning approach

- Metric learning methods have outperformed Bayesian methods and meta-learning methods. For example, ProtoNet.

Shaban, Amirreza, et al. "One-shot learning for semantic segmentation." *arXiv preprint arXiv:1709.03410* (2017).

Problem Definition

Problem Definition

$S=\{(x^i,y^i\}^{Ns}_{i=1}$ denote support set. x^i $[H,W,3]$, y^i represents the corresponding annotation for x^i. Binary mask with $[H,W,1]$

Query image with shape $[H,W,3]$ not in the S .

X may contain multiple semantic classes while the annotation y is only for one class

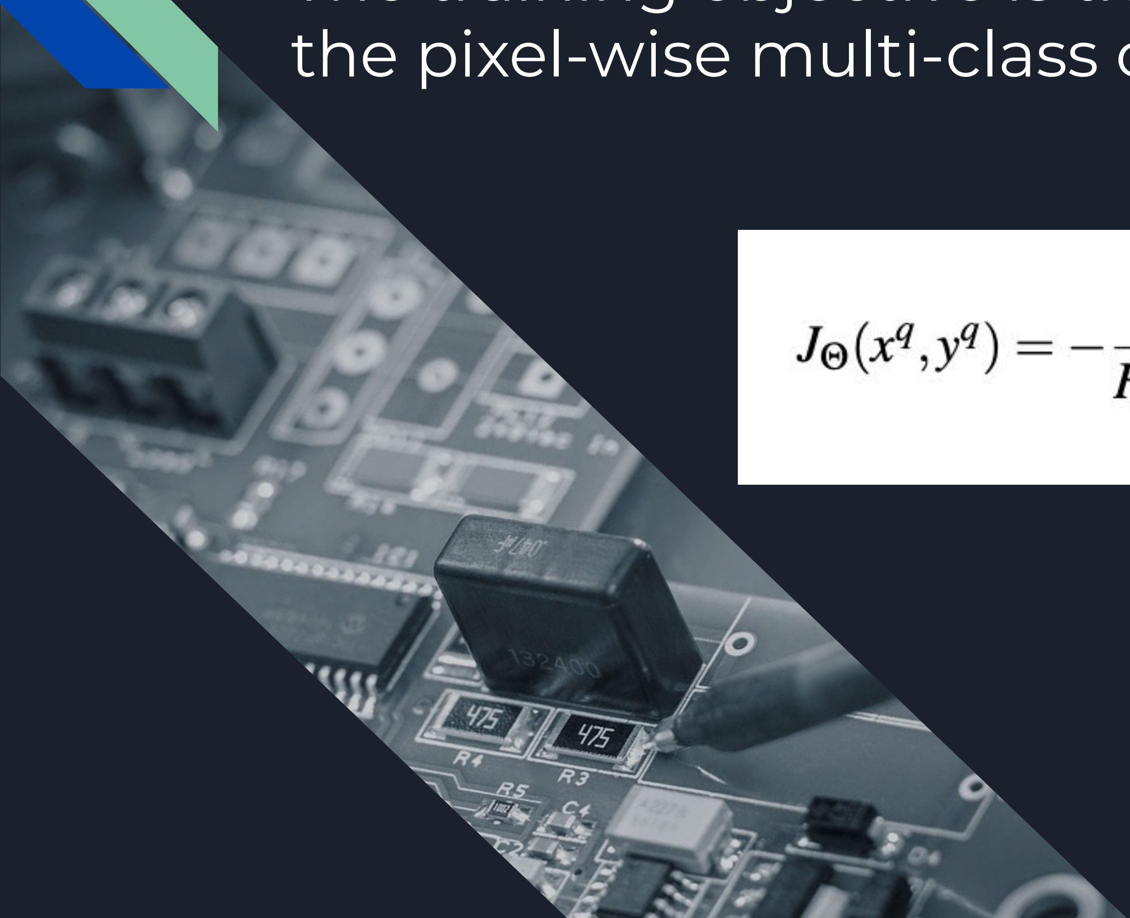
Few shot semantic segmentation N way k shot learning task. Each method is providing with k image mask pairs for each of N classes

Learning $S\text{->} F(-,S)$,

A probability distribution over outputs $F(x^q,S)$

$F_\Theta(x^q,S)$ and the ground truth y^q both have a shape $[H^q, W^q, N+1]$.

The training objective is thus to minimize the pixel-wise multi-class crossentropy loss



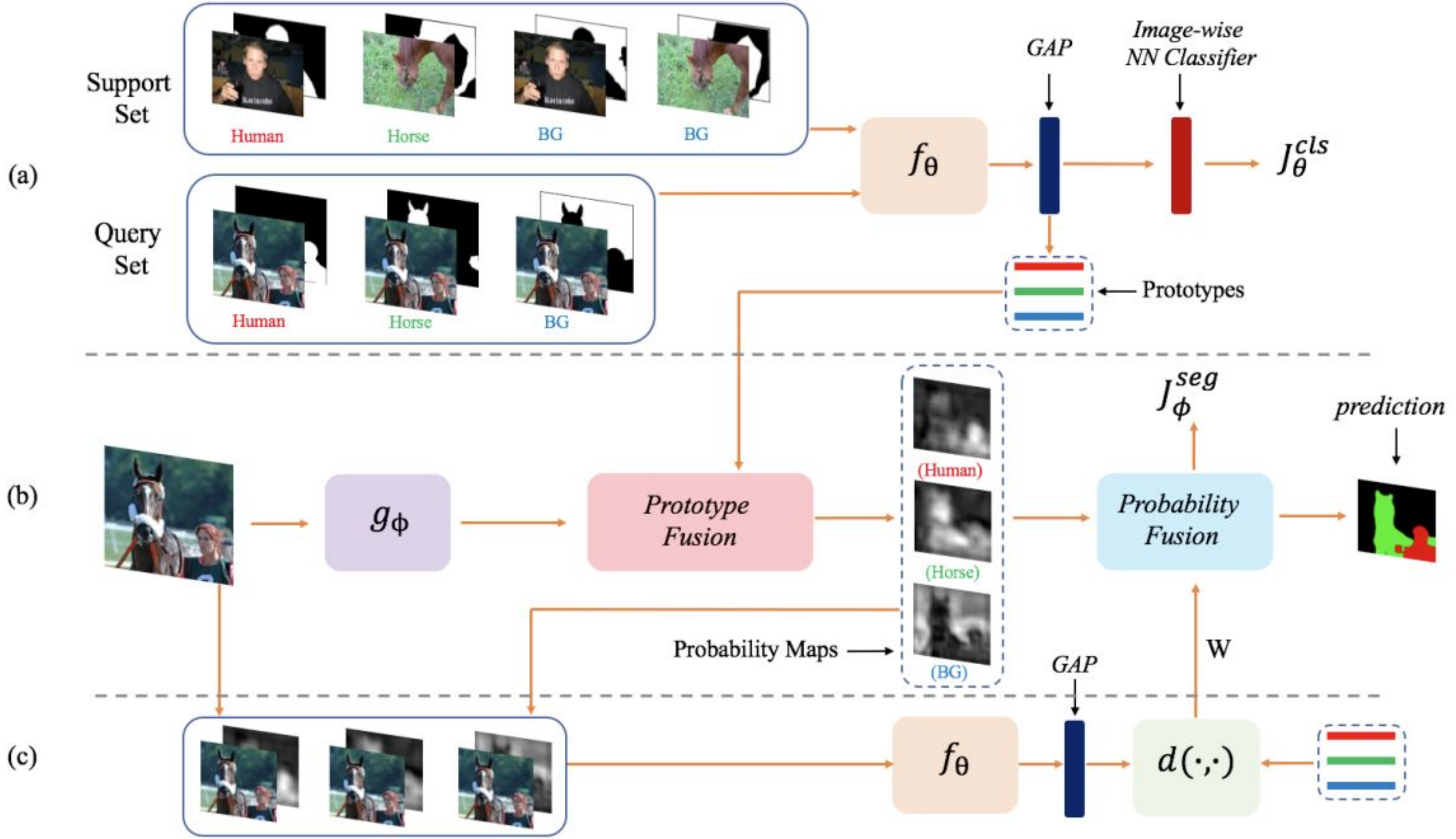
$$J_{\Theta}(x^q, y^q) = -\frac{1}{H^q \times W^q} \sum_{j} \sum_{c} y^q_{j,c} \ln \mathcal{F}_{\Theta}(x^q, S)_{j,c}$$

Few shot Learning Tasks

the test data have classes unseen in the training, these supervised techniques can easily lead to overfitting. We have a seemingly conflicting problem in few-shot semantic segmentation. We want to build a semantic segmentation model but we do not want the model to memorize all the semantic information learned during the training.

Proposed Method

The prototype is a feature vector with highlevel discriminative information. With limited supervision, we train the network in a way that the prediction for a semantic class is close to its prototype in certain projected space



Support Set
Human
Horse
BG
BG
Query Set
Human
Horse
BG
(a)
f_θ
GAP
Image-wise NN Classifier
J_θ^{cls}
Prototypes
(b)
g_ϕ
Prototype Fusion
(Human)
(Horse)
(BG)
Probability Maps
Probability Fusion
J_ϕ^{seg}
prediction
W
(c)
f_θ
GAP
$d(\cdot,\cdot)$

Two Type Loss Function

$$J_{\Theta}(x^q, y^q) = -\frac{1}{H^q \times W^q} \sum_j \sum_c y^q_{j,c} \ln \mathcal{F}_{\Theta}(x^q, S)_{j,c}$$

$$J^{cls}_{\theta} = -\frac{1}{N+1} \sum_t \sum_c I_{c=t} \log(p_{\theta}(y = c | (x^q, y^q_t)))$$

Permutation Training

Algorithm 1: Training an episode for a N-way k-shot semantic segmentation task given a support set S and a query image-label pair (x^q, y^q). A and B are non-zero integers indicating the number of iterations, default value are both 1.

Input: S, (x^q, y^q)
for $a \in \{1, ...A\}$ **do**
  Train the prototype learner by minimizing J_θ^{cls}
end
Get $N+1$ prototypes
Sample B orders from $(N+1)!$ orders uniformly
for $b \in \{1, ...B\}$ **do**
  Order the prototypes
  Train the segmentation model with the ordered prototypes by minimizing J_ϕ^{seg}
end

Experiments

Target audience

mIOU(%)	**1-shot**	**5-shot**
FT	28.1	28.6
Base	34.8	35.0
PL	39.7	40.3
PL + SEG	41.9	42.6
PL + SEG + PT	42.7	43.7

Table 1: Results of 2-way Few-Shot Segmentation on PASCAL-5^i

mIOU(%)	**1-shot**	**5-shot**
FG-BG [3, 26]	55.1	55.6
OSSIS [26, 31]	55.2	-
co-FCN [26]	60.1	60.8
PL	60.0	60.9
PL + SEG	61.2	62.3

Table 2: Results of 1-way Few-Shot Segmentation on PASCAL-5^i

Project objective

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod. Morbi a rutrum nisl. Vestibulum laoreet enim id sem fermentum, sed aliquam arcu dictum. Donec ultrices diam sagittis nibh pellentesque eleifend.

N way Semantic Segmentation

mIOU(%)	**1-shot**	**5-shot**
FT	28.1	28.6
Base	34.8	35.0
PL	39.7	40.3
PL + SEG	41.9	42.6
PL + SEG + PT	42.7	43.7

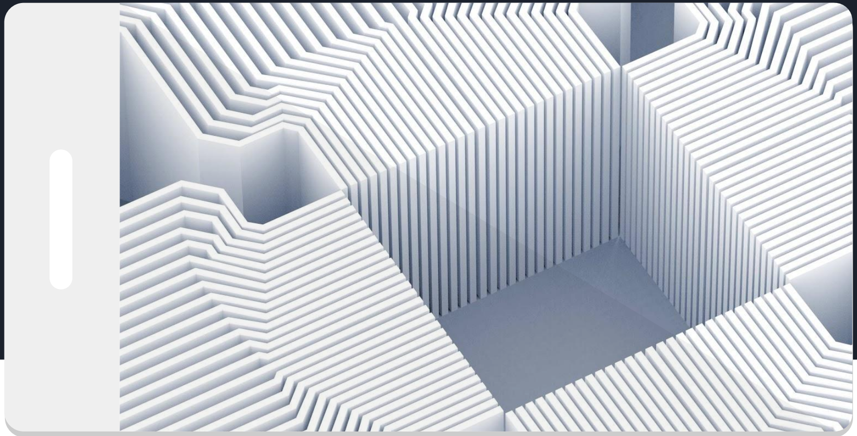
Table 1: Results of 2-way Few-Shot Segmentation on PASCAL-5^i

mIOU(%)	**1-shot**	**5-shot**
FG-BG [3, 26]	55.1	55.6
OSSIS [26, 31]	55.2	-
co-FCN [26]	60.1	60.8
PL	60.0	60.9
PL + SEG	61.2	62.3

Table 2: Results of 1-way Few-Shot Segmentation on PASCAL-5^i

Read Group - Delivered by Yuchen Li

Few-Shot 3D Point Cloud Semantic Segmentation



Zhao, Na, Tat-Seng Chua, and Gim Hee Lee. "Few-shot 3d point cloud semantic segmentation." *arXiv preprint arXiv:2006.12052* (2020).

Few-shot 3D Point Cloud Semantic Segmentation

Introduction

Related Work

Our Methodology

Experiments

Conclusion

Overview

Existing point cloud semantic segmentation methods require a large number of tags, it is difficult to mark point clouds point by point, and it has poor generalization for new categories. Therefore, this paper proposes a significantly-perceived few-shot segmentation method for multi-prototype transformation.

Specifically, each class can be represented as multiple prototypes to model complex distributed location cloud data. Then, a label propagation method is designed to explore the multi-prototype affinity between labeled and unlabeled points, and the affinity between unlabeled points. Finally, the significant perception of multi-scale features learns the semantic and geometric correlations between network modeling points.

Few-shot point cloud semantic segmentation task

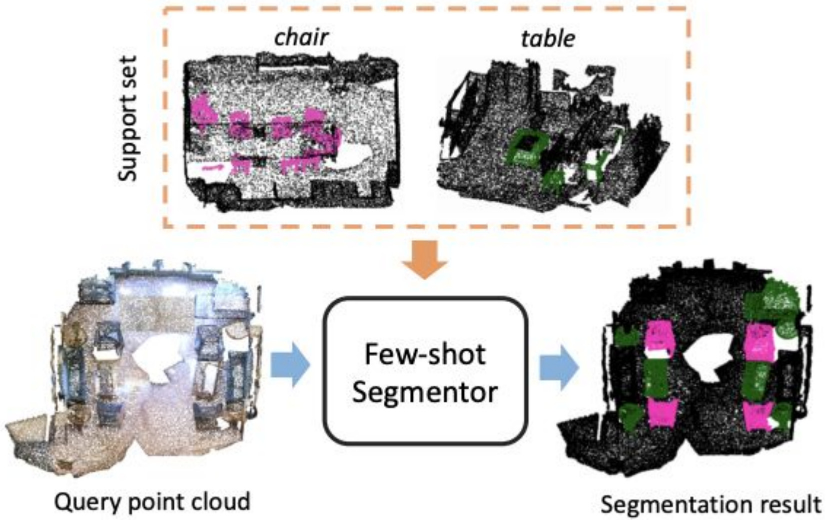


Figure 1. Few-shot point cloud semantic segmentation task is to learn a segmentor that segments the query point cloud in terms of new classes with learned knowledge from the support examples. This figure illustrates an example with 2-way 1-shot setting.

Overview

This paper propose a novel attention-aware multi prototype transductive inference method for few-shot point cloud semantic segmentation

Model complex distributions of the points within the point clouds of the support set, and perform the segmentation iva transductive inference with the discriminative features extracted under the few-shot constraint.

Attention-aware Multi-prototype Transductive Inference Method

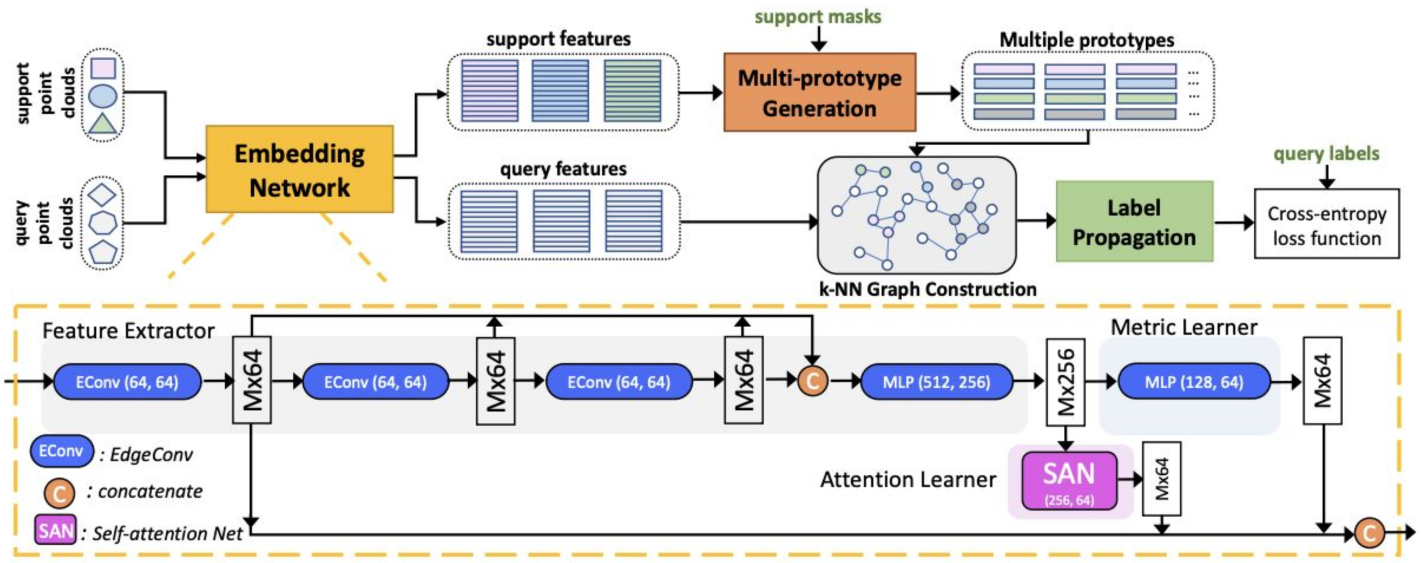


Figure 2. The architecture of our proposed method. This figure illustrates a 3-way 1-shot setting.

Embedding Network

1> encode the geometric structures of points based on local context

2> encode the semantic information of points and their semantic correlation based on global context

3> quickly adopt to different few shot tasks

Feature extractor, attention learner, and metric learner.

Self attention network

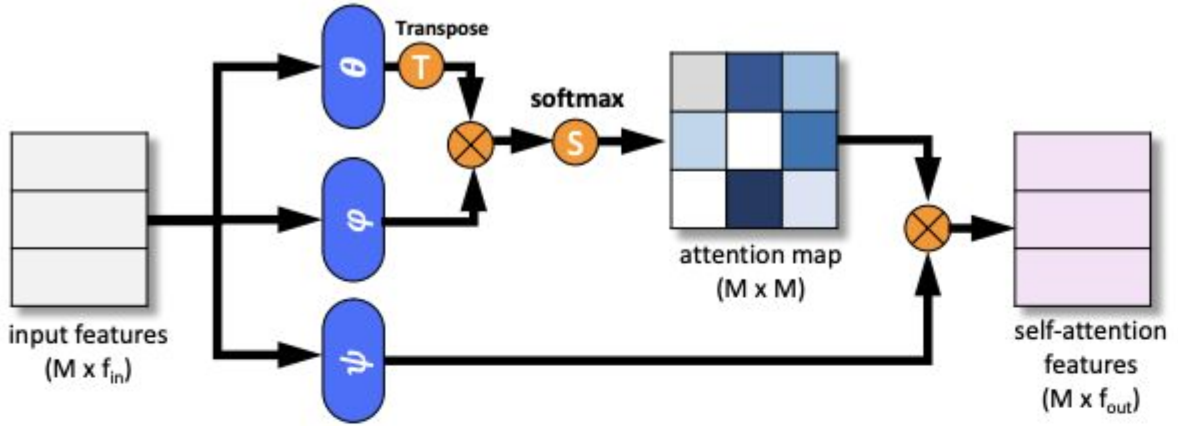


Figure 3. The architecture of Self Attention Network (SAN). θ, φ, and ψ are linear embedding functions with trainable parameters.

Multi-prototype Generation

Sampling seed point and point-to-seed assignment based on the learned embedding space.

$$\boldsymbol{\mu}^c = \left\{ \boldsymbol{\mu}_1^c, \ldots, \boldsymbol{\mu}_n^c \mid \boldsymbol{\mu}_i^c = \frac{1}{|\mathcal{I}_i^{c*}|} \sum_{\mathbf{f}_j^c \in \mathcal{I}_i^{c*}} \mathbf{f}_j^c \right\}$$

$$\text{s.t.} \quad \mathcal{I}^{c*} = \underset{\mathcal{I}^c}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{\mathbf{f}_j^c \in \mathcal{I}_i^c} \|\mathbf{f}_j^c - \mathbf{s}_i^c\|_2, \tag{2}$$

where $\{\mathbf{f}_i^c\}_{i=1}^{m^c}$ is partition into n sets $\mathcal{I}^{c*} = \{\mathcal{I}_1^{c*}, \ldots, \mathcal{I}_n^{c*}\}$ such that $\mathbf{f}_j^c \in \mathcal{I}_i^{c*}$ is assigned to $\mathbf{s}_i^c$.

Transductive Inference

K NN graph construction

n* (N+1) multiprototypes and T * M query points as nodes of a graph with size V=n * (N+1) +T *M. Construct a sparse affinity matrix, by computing the Gaussian similarity between each node and its k nearest neighbors in the embedding space:

$$\mathbf{A}_{ij} = \exp\left(-\frac{||\mathbf{v}_i - \mathbf{v}_j||_2^2}{2\sigma^2}\right),\ for\ \mathbf{v}_j \in \mathcal{N}_k(\mathbf{v}_i), \qquad (3)$$

W=A+A^T, symmetrically normalize W to yield S=D^-1/2WD^-1/2

$\mathbf{Y} \in \mathbb{R}^{V\times(N+1)}$,

Label Propagation

$$\mathbf{Z}_{t+1} = \alpha \mathbf{S} \mathbf{Z}_t + (1 - \alpha)\mathbf{Y}.$$

Sequence {Zt} converges to a closed-form solution:

$$\mathbf{Z}^* = (\mathbf{I} - \alpha \mathbf{S})^{-1}\mathbf{Y}.$$

Experiments Result

Method	2-way						3-way					
	1-shot			5-shot			1-shot			5-shot		
	S^0	S^1	mean	S^0	S^1	mean	S^0	S^1	mean	S^0	S^1	mean
FT	36.34	38.79	37.57	56.49	56.99	56.74	30.05	32.19	31.12	46.88	47.57	47.23
ProtoNet	48.39	49.98	49.19	57.34	63.22	60.28	40.81	45.07	42.94	49.05	53.42	51.24
AttProtoNet	50.98	51.90	51.44	61.02	65.25	63.14	42.16	46.76	44.46	52.20	56.20	54.20
MPTI	52.27	51.48	51.88	58.93	60.56	59.75	44.27	46.92	45.60	51.74	48.57	50.16
Ours	**53.77**	**55.94**	**54.86**	**61.67**	**67.02**	**64.35**	**45.18**	**49.27**	**47.23**	**54.92**	**56.79**	**55.86**

Table 1: Results on **S3DIS** dataset using mean-IoU metric (%). S^i denotes the split i is used for testing.

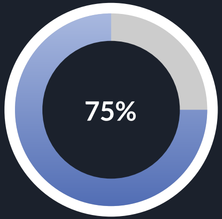
Method	2-way						3-way					
	1-shot			5-shot			1-shot			5-shot		
	S^0	S^1	mean	S^0	S^1	mean	S^0	S^1	mean	S^0	S^1	mean
FT	31.55	28.94	30.25	42.71	37.24	39.98	23.99	19.10	21.55	34.93	28.10	31.52
ProtoNet	33.92	30.95	32.44	45.34	42.01	43.68	28.47	26.13	27.30	37.36	34.98	36.17
AttProtoNet	37.99	34.67	36.33	52.18	46.89	49.54	32.08	28.96	30.52	44.49	39.45	41.97
MPTI	39.27	36.14	37.71	46.90	43.59	45.25	29.96	27.26	28.61	38.14	34.36	36.25
Ours	**42.55**	**40.83**	**41.69**	**54.00**	**50.32**	**52.16**	**35.23**	**30.72**	**32.98**	**46.74**	**40.80**	**43.77**

Table 2: Results on **ScanNet** dataset using mean-IoU metric (%). S^i denotes the split i is used for testing.

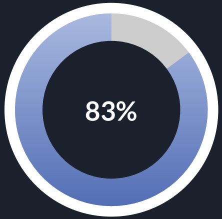
Persona 01

Wendy Writer

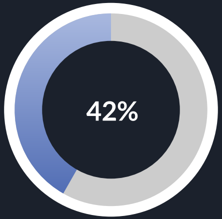
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.



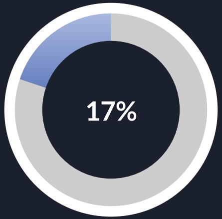
Lorem Ipsum



Lorem Ipsum



Lorem Ipsum

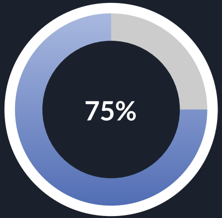


Lorem Ipsum

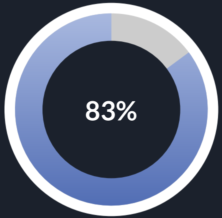
Persona 02

Berry Books

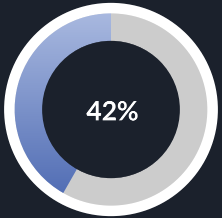
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.



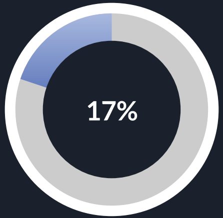
Lorem Ipsum



Lorem Ipsum



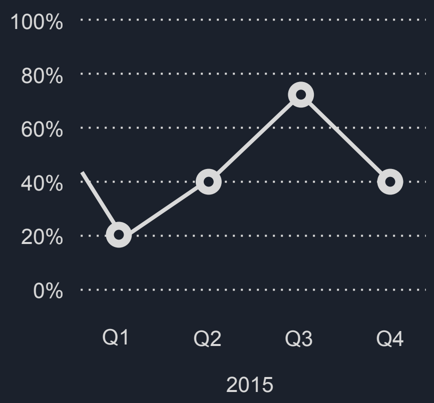
Lorem Ipsum

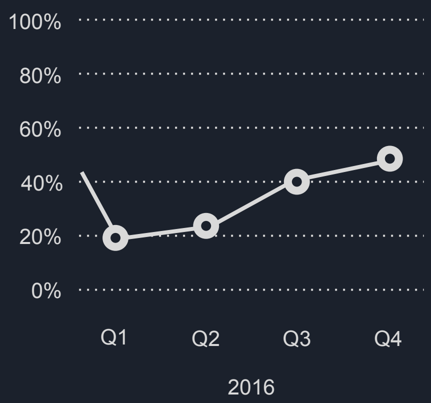


Lorem Ipsum

Market trends

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod. Morbi a rutrum nisl. Vestibulum laoreet enim id sem fermentum, sed aliquam arcu dictum. Donec ultrices diam sagittis nibh pellentesque eleifend.





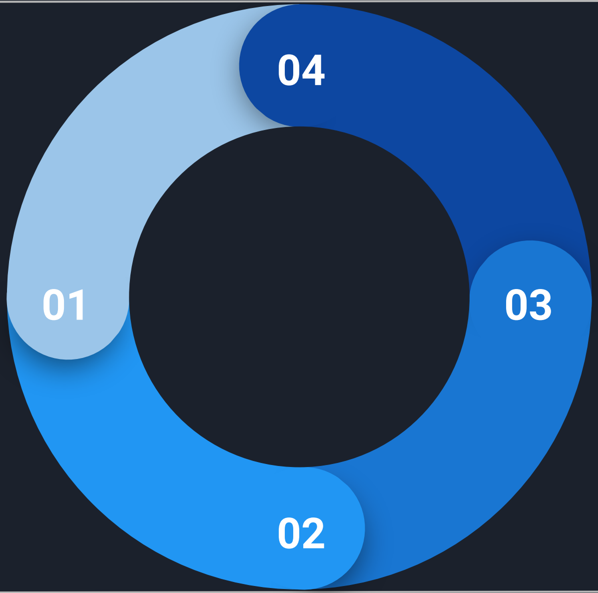
Cycle diagram

Prototype

Lorem ipsum dolor sit amet, consectetur adipiscing.

Share

Lorem ipsum dolor sit amet, consectetur adipiscing.



Refine

Lorem ipsum dolor sit amet, consectetur adipiscing.

Get feedback

Lorem ipsum dolor sit amet, consectetur adipiscing.

Introducing: Lorem ipsum

Showcase how your tools work across different devices

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum.

QUICK TIP

Try right clicking on a photo and using "Replace Image" to demonstrate your prototype.

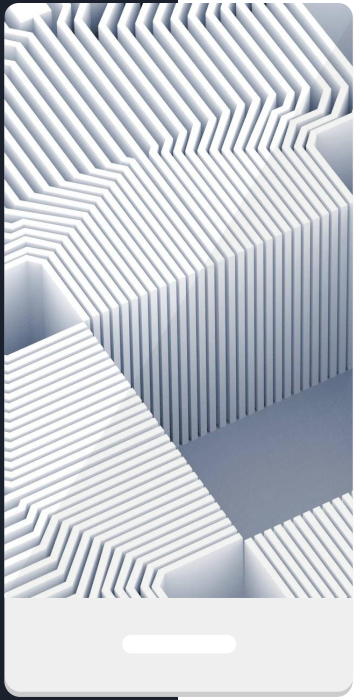
Spotlight on desktop

Lorem ipsum dolor sit consectetur amet adipiscing donec

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.

Spotlight on mobile

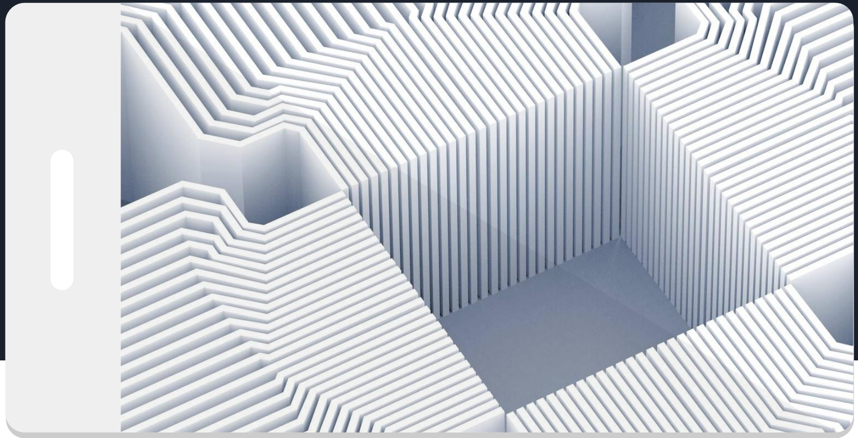
Lorem ipsum dolor sit consectetur amet adipiscing donec



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.

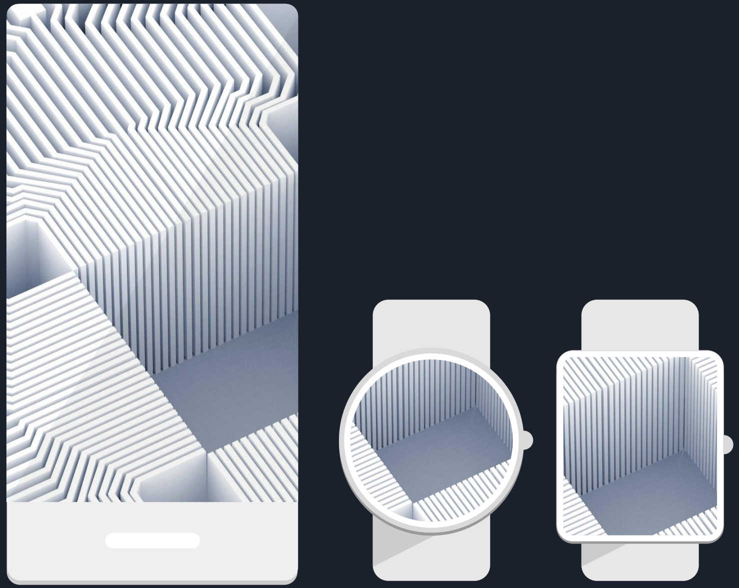
Spotlight on landscape view on mobile

Lorem ipsum dolor sit consectetur amet adipiscing donec



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.

Spotlight on wearables

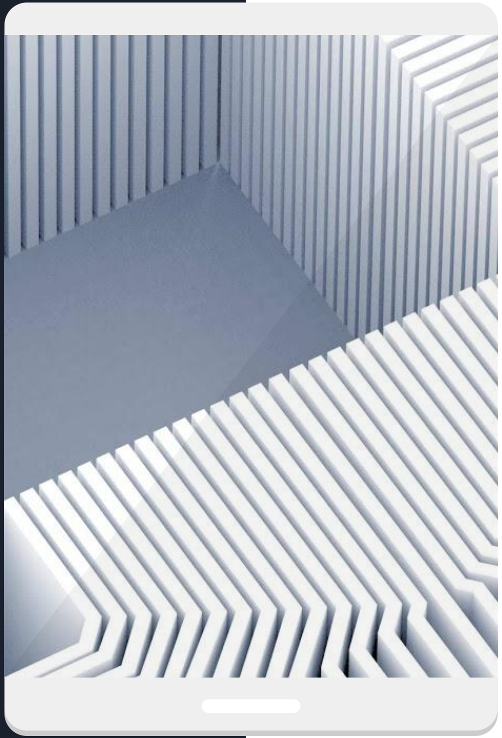


Lorem ipsum dolor sit consectetur amet adipiscing donec

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.

Spotlight on tablet

Lorem ipsum dolor sit consectetur amet adipiscing donec

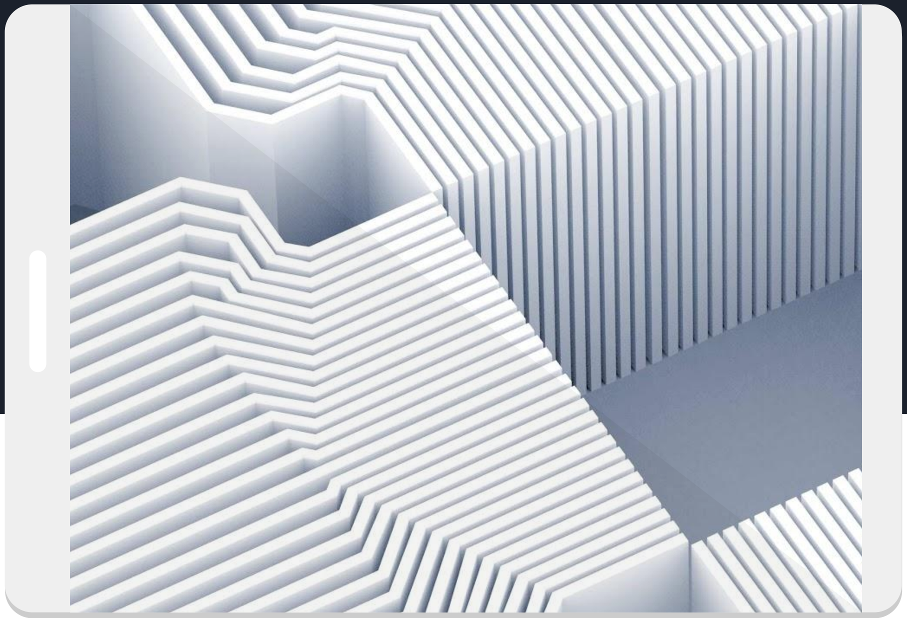


Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.

Spotlight on landscape view on tablet

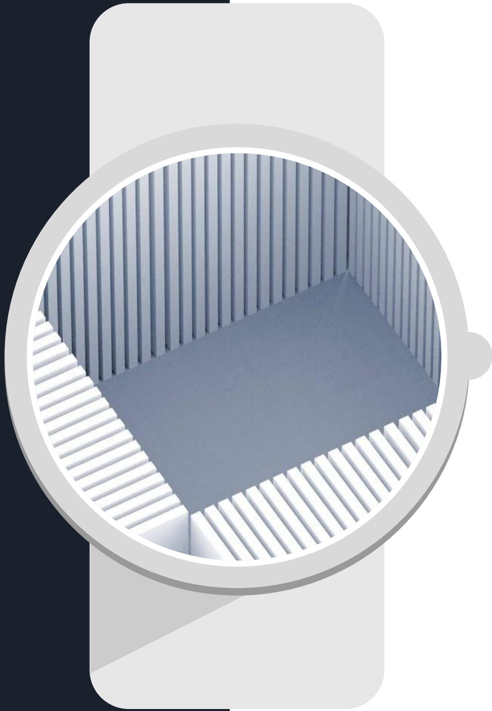
Lorem ipsum dolor sit consectetur amet adipiscing donec

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.



Spotlight on wearables

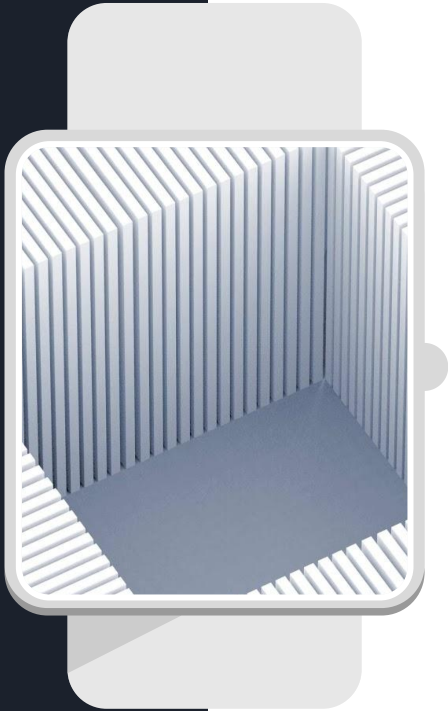
Lorem ipsum dolor sit consectetur amet adipiscing donec



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.

Spotlight on wearables

Lorem ipsum dolor sit consectetur amet adipiscing donec



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit. Fusce venenatis nunc ut lectus convallis, sit amet egestas mi rutrum. Maecenas molestie ultricies euismod.

Project timeline

JAN	FEB	MAR	APR	MAY	JUN
Lorem Ipsum	**Sit Amet**	**Donec Ultrices**	**Litora**	**Quis Diam**	**Dolor Lorem**
Lorem ipsum dolor sit amet, consectetur adipiscing.	Lorem ipsum dolor sit amet, consectetur adipiscing.	Lorem ipsum dolor sit amet, consectetur adipiscing.	Lorem ipsum dolor sit amet, consectetur adipiscing.	Lorem ipsum dolor sit amet, consectetur adipiscing.	Lorem ipsum dolor sit amet, consectetur adipiscing.

Thank you!

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eleifend a diam quis suscipit.