Plane Geometry Diagram Parsing

Ming-Liang Zhang[1,2,*], Fei Yin[1,2], Yi-Han Hao[1,3], Cheng-Lin Liu[1,2]

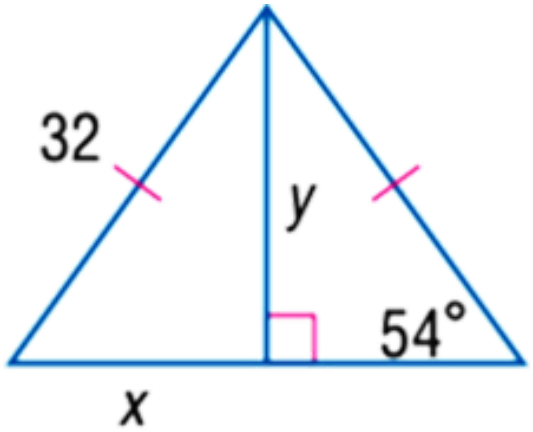
[1]National Laboratory of Pattern Recognition, Institute of Automation of Chinese Academy of Sciences

[2]School of Artificial Intelligence, University of Chinese Academy of Sciences

[3]School of Electronic Information Engineering, Beijing Jiaotong University

Github: https://github.com/mingliangzhang2018/PGDP

Task Introduction



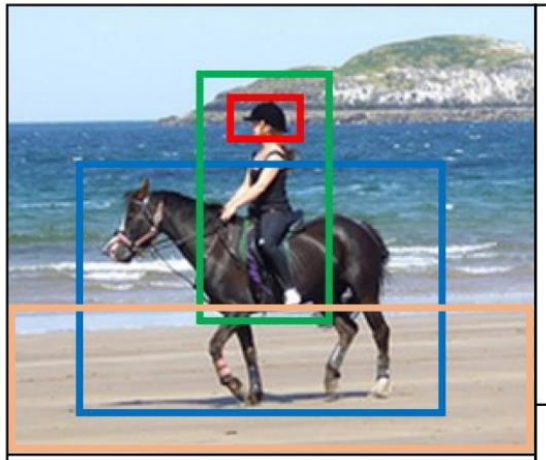
Problem Text:
Find x. Round to the nearest tenth.
Choices:
A.18.8 B.23.2 C.25.9 D.44.0
Answer: A


```
"diagram_logic_form": [
  "Equals(MeasureOf(Angle(D, C, B)), 54)",
  "Equals(LengthOf(Line(A, D)), x)",
  "Equals(LengthOf(Line(B, D)), 32)",
  "Equals(LengthOf(Line(A, B)), y)",
  "PointLiesOnLine(A, Line(C, D))",
  "Equals(LengthOf(Line(B, D)), LengthOf(Line(B, C)))",
  "Perpendicular(Line(A, B), Line(A, C))"
],
"line_instances": [
  "AB", "AC", "AD", "BC", "BD", "CD"
],
"circle_instances": [],
"point_positions": {
  "A": [
    101.98, 137.25
  ],
  "B": [
    100.56, 1.07
  ],
  "C": [
    201.05, 137.03
  ],
  "D": [
    0.82, 136.66
  ]
```

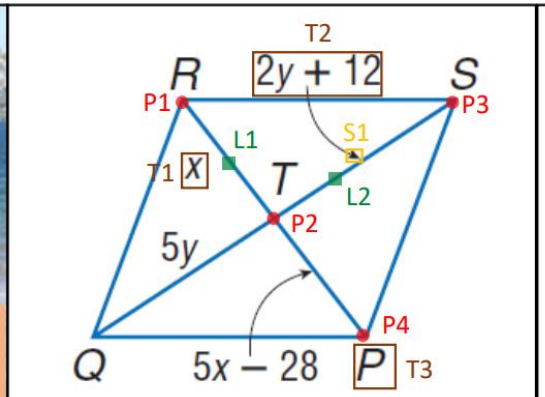

- **Task 1:**
  Extract geometric primitives (point, line and circle)
- **Task 2:**
  Detect non-geometric primitives (symbol and text)
- **Task 3:**
  Build basic relationships among primitives
- **Task 4:**
  Generate the geometry formal language

Task Modeling

Compare SGG with PGDP

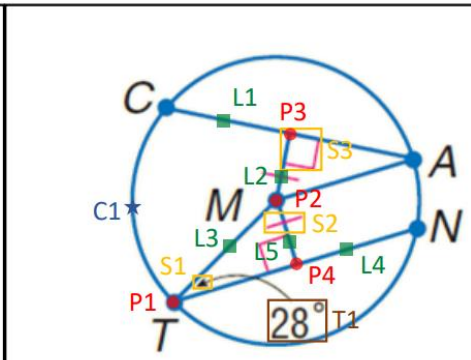


- (girl, ride, horse)
- (girl, wear, hat)
- (horse, on, beach)



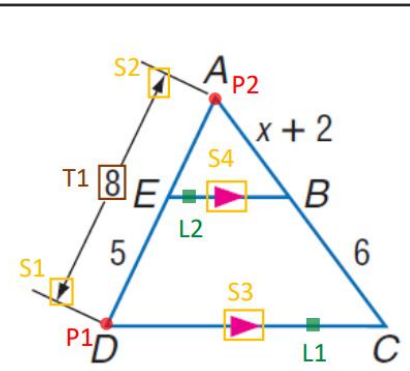
- (T1, (P1, P2, L1))
- (T2, (S1))
- (T3, (P4))
- (S1, (P2, P3, L2))
- (P2, (L1))
- (P2, (L2))

.....



- (T1, (S1))
- (S1, (P1, L3 , L4))
- (P1, (C1))
- (P2, (C1))
- (S3, (P3, L1 , L2))
- (S2, (P2, P4 , L5))

.....



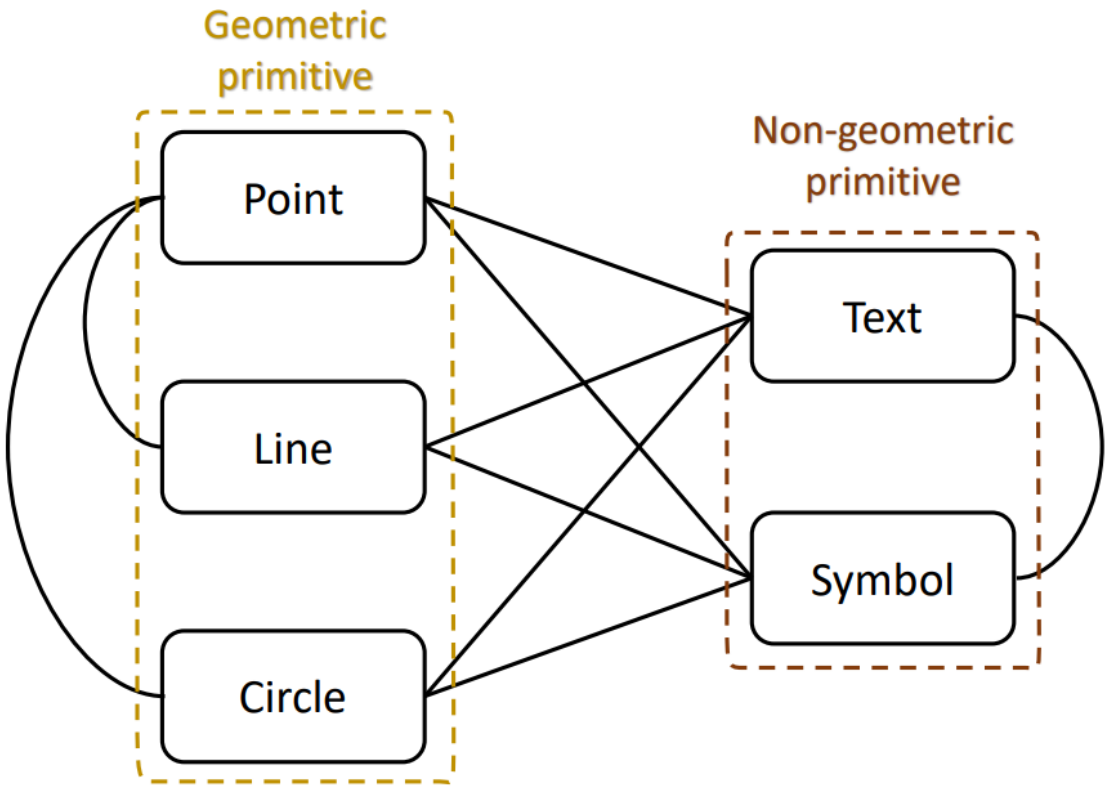
- (T1, (S1 , S2))
- (S1, (P1))
- (S2, (P2))
- (S3, (L1))
- (S4, (L2))

.....

- **SGG (scene graph generation) :**
  - ✓ Coarse-grained box position
  - ✓ Subject-predicate-object triplet
  - ✓ Explicit relationship prediction

- **PGDP (plane geometry diagram parsing) :**
  - ✓ Fine-grained instance mask
  - ✓ Two-tuple with multiple entities
  - ✓ No need of relationship classification

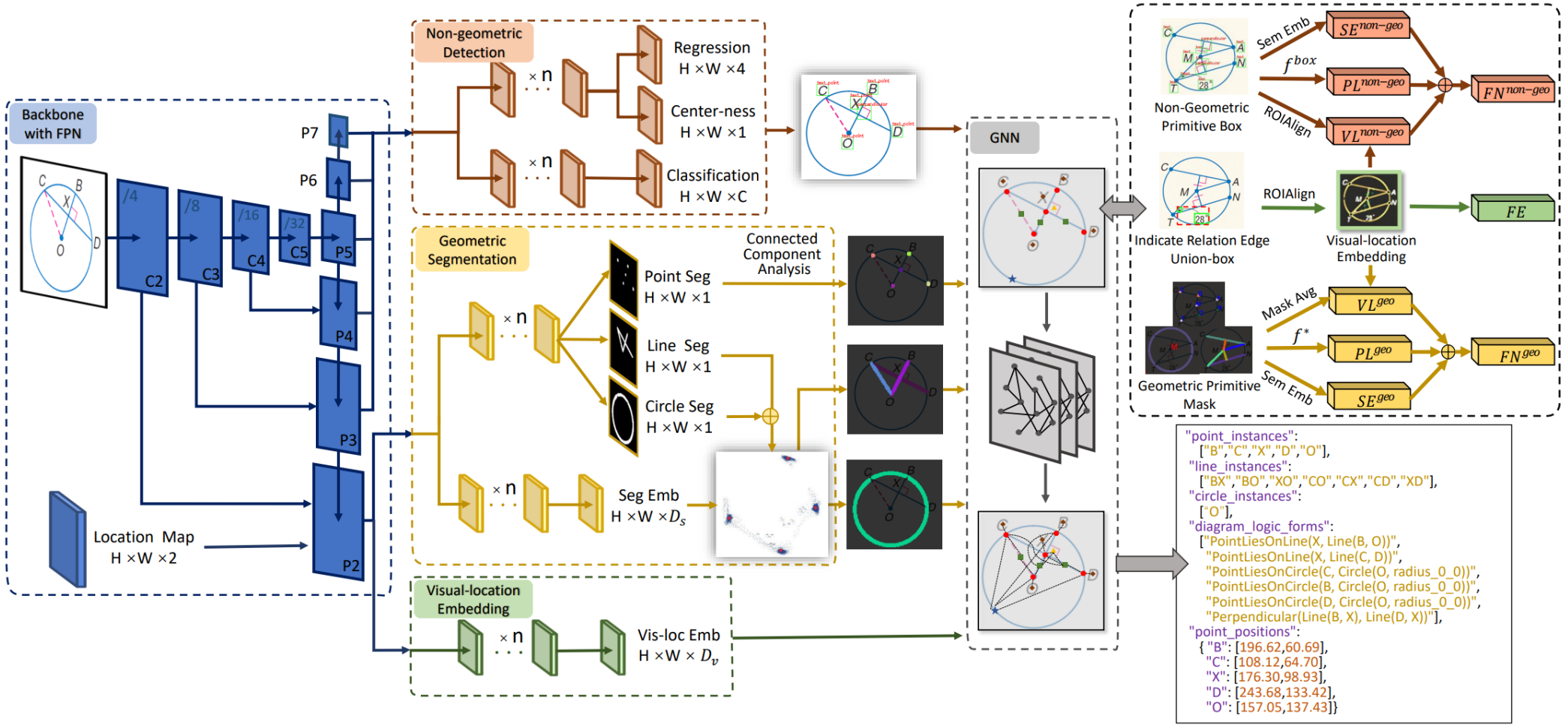
Task Modeling



Graph of relationships among primitives

- Only construct relations between point and line, point and circle for relations among geometric primitives
- Take point, symbol and text as subjects, and serve other related primitives as objects, neglecting the predicate term
- Served as relation annotation manner of dataset and composition method of GNN model

PGDPNet Model

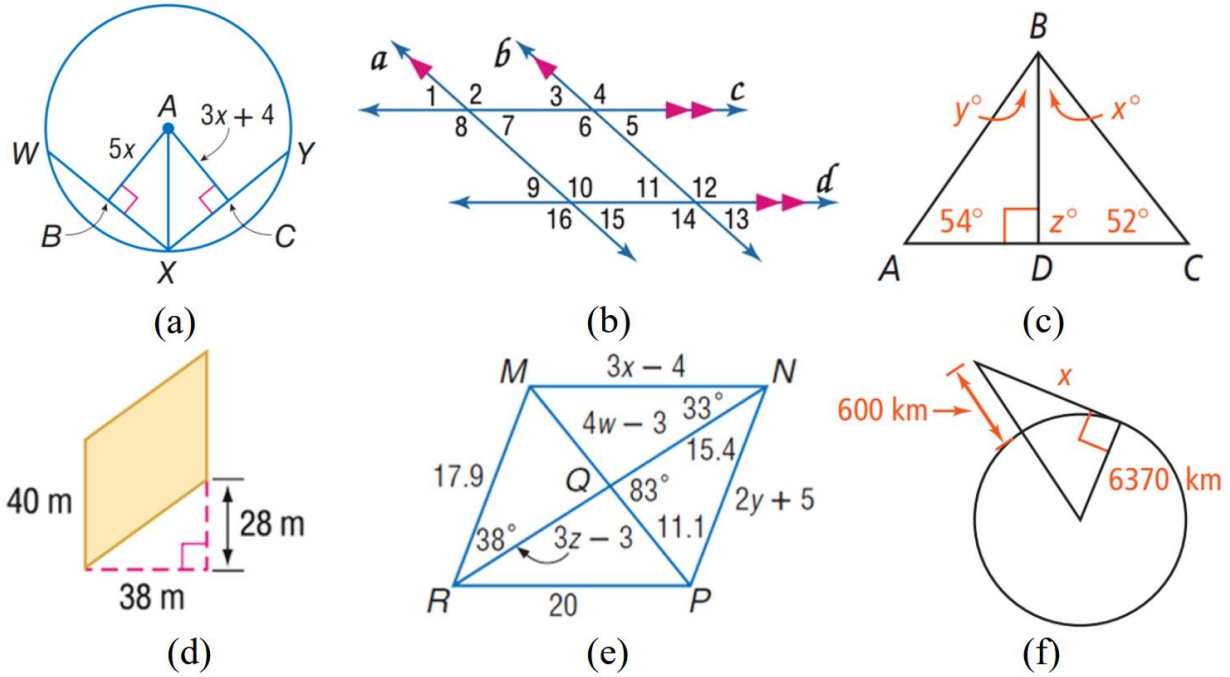


PGDPNet consists of five modules: Backbone module (**BM**), Non-geometric Detection module (**NDM**), Geometric Segmentation module (**GSM**), Visual-location Embedding module (**VLEM**), GNN module (**GM**)

PGDPNet Model

- NDM detects non-geometric primitives with FCOS heads.
- GSM conducts instance segmentation of geometric primitives at pixel level.
- GM parses relationships among primitives by edges predication and fine-grained text classification.
- PGDPNet realizes multi-task learning in end-to-end style with weighted sum losses of NDM, GSM and GM.
- Geometry formal language is generated according to sub-tasks results, geometry prior knowledge and language grammar.

PGDP5K Dataset



Plane geometry diagram with various styles, primitive interference, complex layouts in PGDP5K dataset

- Consists of 5,000 samples, where 1,813 samples are selected from Geometry3K dataset and other 3,187 samples are collected from three textbooks across grades 6-12
- labeled with fine-grained annotations at primitive level, including primitive classes, locations and relationships
- Generates geometric propositions by combining with above annotations and geometry prior knowledge

Task Performance

Geometric Primitive Detection

		Freeman	GEOS	PGDPNet
Point	Precision	67.46	76.51	**99.65**
	Recall	80.41	93.44	**99.71**
	F1	73.37	84.13	**99.68**
Line	Precision	50.78	66.99	**99.30**
	Recall	80.43	90.46	**99.51**
	F1	62.25	76.98	**99.40**
Circle	Precision	90.72	98.25	**99.85**
	Recall	97.75	99.24	**99.96**
	F1	94.10	98.74	**99.90**

Primitive Relation Parsing

		Baseline	w SE	w PL	w SE&PL
All	Precision	98.96	99.16	99.10	**99.16**
	Recall	96.97	97.01	**97.11**	97.07
	F1	97.96	98.07	98.08	**98.11**
Geo2Geo	Precision	98.84	**99.15**	99.09	99.13
	Recall	98.53	98.56	**98.69**	98.60
	F1	98.68	98.85	**98.89**	98.86
Text2Geo	Precision	99.09	99.27	**99.38**	99.27
	Recall	96.16	96.61	96.36	**96.84**
	F1	97.60	97.94	97.84	**98.04**
Sym2Geo	Precision	99.06	98.71	99.01	**99.07**
	Recall	94.13	94.89	95.02	**95.27**
	F1	96.53	96.76	96.97	**97.13**
Text2Head	Precision	97.70	98.03	98.03	**98.08**
	Recall	91.95	92.26	92.57	**95.05**
	F1	94.74	95.06	95.22	**96.54**
Complete Acc		81.50	82.50	82.60	**83.20**

Task Performance

Geometry Formal Language Generation

		IMP-Geometry3K	PGDP5K
All	Likely Same	73.71 / 99.17 / **99.33**	65.70 / 98.40 / **99.00**
	Almost Same	50.08 / 95.51 / **98.50**	44.40 / 93.10 / **96.60**
	Perfect Recall	45.26 / 81.03 / **92.18**	40.00 / 79.70 / **86.20**
	Totally Same	34.28 / 80.53 / **91.51**	27.30 / 78.20 / **84.70**
Geo2Geo	Likely Same	69.88 / **99.67** / 99.50	63.90 / **99.10** / 99.00
	Almost Same	56.24 / **99.50** / 99.00	49.40 / **97.30** / 97.10
	Perfect Recall	74.71 / **99.33** / 99.17	78.70 / 96.90 / **97.40**
	Totally Same	47.59 / **98.84** / 98.33	40.80 / 93.60 / **94.50**
Non-geo 2Geo	Likely Same	77.04 / 96.01 / **99.00**	67.30 / 95.80 / **98.00**
	Almost Same	59.07 / 89.35 / **96.01**	49.80 / 88.20 / **94.90**
	Perfect Recall	50.92 / 81.20 / **92.85**	45.70 / 81.30 / **87.00**
	Totally Same	48.59 / 80.87 / **92.85**	40.50 / 80.60 / **86.40**

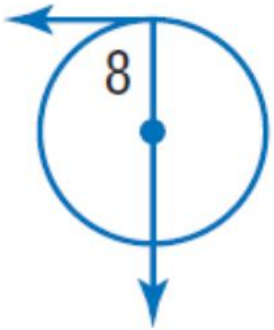
"&/&/&" denotes performances of three methods InterGPS, PGDPNet without GNN, PGDPNet.

Geometry Problem Solving

	Text InterGPS	Text GT
Diagram w/o	25.4±0.0	25.4±0.0
Diagram InterGPS	57.5±0.2	58.0±1.7
Diagram PGDPNet w/o GNN	69.3±0.2	70.0±0.4
Diagram PGDPNet	74.1±0.2	74.3±0.3
Diagram GT	**75.9±0.2**	**76.0±0.4**

Task Performance

Failure Cases of PGDPNet



Problem Text:
Find angle 8.

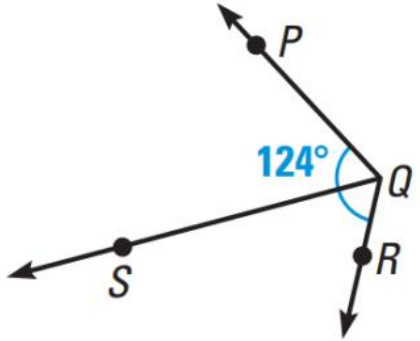
Choices:
A. 20 B. 70 C. 90 D. 180

Answer: C

(a)

Error:
Equals(LengthOf(Line(A, B)), 8)

True:
Equals(MeasureOf(Angle(A, B, C)), MeasureOf(angle 8))



Problem Text:
$\overrightarrow{QS}$ is angle bisector of ∠PQR, what are the measures of ∠PQS and ∠SQR?

Answer: 62°

(b)

Error:
Equals(MeasureOf(Angle(P, Q, S)), 124)

True:
Equals(MeasureOf(Angle(P, Q, R)), 124)

Conclusion

- We propose PGDPNet, the first end-to-end deep learning model for explicit geometry diagram parsing.
- We build a large-scale dataset PGDP5K, containing fine-grained annotations of primitives and relations.
- Our method demonstrates superior performance of geometry diagram parsing, outperforming previous methods significantly.
- Future works will consider incorporating parsing content of diagram and problem text to improve performance of geometry problem solving.

Thanks for Your Listening!

Code Link

Paper Link

Dataset Link

Contact us: zhangmingliang2018@ia.ac.cn