Publications

Journal Publication

- **Shreya Goyal**, Naimul Khan, Chiranjoy Chattopadhyay, Gaurav Bhatnagar, "GRIHA: Synthesizing 2-Dimensional Building Layouts from Images Captured using a Smart Phone", Multimedia Tools and Applications, January 2022.

- **Shreya Goyal**, Chiranjoy Chattopadhyay, Gaurav Bhatnagar,"Knowledge driven Description Synthesis for Floor Plan Interpretation", International Journal on Document Analysis and Recognition (IJDAR), April 2021, DOI:10.1007/s10032-021-00367-3

- **Shreya Goyal**, Satya Bhavsar, Shreya Patel, Chiranjoy Chattopadhyay,Gaurav Bhatnagar; "SUGAMAN:Describing Floor Plans for Visually Impaired by Annotation Learning and Proximity based Sentence Model" IET Image Processing, 2019, 13, (13), p. 2623-2635, DOI: 10.1049/iet-ipr.2018.5627

Conference Publications

- **Shreya Goyal**, Naimul Khan, Chiranjoy Chattopadhyay, Gaurav Bhatnagar; "LayART: Generating indoor layout using ARCore Transformations", IEEE International Conference on Multimedia Big Data (BigMM), September 2020., p. 272–276, DOI: 10.1109/BigMM50055.2020.00047

- **Shreya Goyal**, Vishesh Mistry, Chiranjoy Chattopadhyay, Gaurav Bhatnagar: "BRIDGE: Building plan Repository for Image Description Generation, and Evaluation", IAPR International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia September, 2019, p. 1071–1076, DOI:10.1109/ICDAR.2019.00174.

- Mahak Jain, Anurag Sanyal, **Shreya Goyal**, Chiranjoy Chattopadhyay,Gaurav Bhatnagar; "A Framework for the Conversion of Textual Big Data into 2D Architectural Floor Plan", 1st International Workshop on Big Data in Culture, Design and Heritage, IEEE BigMM, Singapore, September 2019, p. 404–410, DOI: 10.1109/BigMM.2019.00022

- **Shreya Goyal**, Chiranjoy Chattopadhyay , Gaurav Bhatnagar; "ASYSST: A Framework for Synopsis Synthesis Empowering Visually Impaired", Workshop on Multimedia for Accessible Human Computer Interface, ACM Multimedia (MM 2018), Seoul, Korea, October 2018, p. 17–24, DOI: 10.1145/3264856.3264859

- **Shreya Goyal**, Chiranjoy Chattopadhyay, Gaurav Bhatnagar: "Plan2Text: A Framework for Describing Building Floor Plan Images From First Person Perspective floor plan images", IEEE Colloquiumon Signal Processing and its Applications,(CSPA 2018), Malayasia, March 2018, p. 35–40 , DOI: 10.1109/CSPA.2018.8368681

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S., "SLIC superpixels compared to state-of-the-art superpixel methods", *Transaction on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, 2274–2282 2012

Ah-Soon, C., "A constraint network for symbol detection in architectural drawings", in "International Workshop on Graphics Recognition", pp. 80–90, Springer 1997

Ah-Soon, C. and Tombre, K., "Variations on the analysis of architectural drawings", in "International Conference on Document Analysis and Recognition", vol. 1, pp. 347–351, IEEE 1997

Ahmed, S., Liwicki, M., Weber, M., and Dengel, A., "Improved automatic analysis of architectural floor plans", in "International Conference on Document Analysis and Recognition", 2011

Ahmed, S., Liwicki, M., Weber, M., and Dengel, A., "Automatic room detection and room labeling from architectural floor plans", in "International Workshop on Document Analysis Systems", pp. 339–343, IEEE 2012

Alhashim, I. and Wonka, P., "High Quality Monocular Depth Estimation via Transfer Learning", *arXiv preprint arXiv:1812.11941* 2018

Almazán, J., Gordo, A., Fornés, A., and Valveny, E., "Word spotting and recognition with embedded attributes", *Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 12, 2552–2566 2014

Anderson, P., Fernando, B., Johnson, M., and Gould, S., "Spice: Semantic propositional image caption evaluation", in "European Conference on Computer Vision", pp. 382–398, Springer 2016

Angladon, V., *Room layout estimation on mobile devices*, Ph.D. thesis, University of Toulouse 2018

Aoki, Y., Shio, A., Arai, H., and Odaka, K., "A prototype system for interpreting hand-sketched floor plans", in "International Conference on Pattern Recognition", vol. 3, pp. 747–751, IEEE 1996

Aptoide, "AR Plan 3D Ruler – Camera to Plan, Floor planner", `https://arplan-3d.en.aptoide.com/` 2020

Arduengo, M., Torras, C., and Sentis, L., "Robust and Adaptive Door Operation with a Mobile Manipulator Robot", *arXiv preprint arXiv:1902.09051* 2019

Bahdanau, D., Cho, K., and Bengio, Y., "Neural machine translation by jointly learning to align and translate", *arXiv preprint arXiv:1409.0473* 2014

Baird, H. S., "Digital libraries and document image analysis", in "International Conference on Document Analysis and Recognition", pp. 2–14, IEEE 2003

Banerjee, S. and Lavie, A., "METEOR: An automatic metric for MT evaluation with improved correlation with human judgments", in "Workshop on Statistical Machine Translation", 2005

Bao, S. Y., Furlan, A., Fei-Fei, L., and Savarese, S., "Understanding the 3D layout of a cluttered room from multiple images", in "Winter Conference on Applications of Computer Vision", pp. 690–697, IEEE 2014

Barducci, A. and Marinai, S., "Object recognition in floor plans by graphs of white connected components", in "International Conference on Pattern Recognition", pp. 298–301, IEEE 2012

Baró, A., Riba, P., Calvo-Zaragoza, J., and Fornés, A., "From optical music recognition to handwritten music recognition: A baseline", *Pattern Recognition Letters*, vol. 123, 1–8 2019

Bernardi, R., Cakici, R., Elliott, D., Erdem, A., Erdem, E., Ikizler-Cinbis, N., Keller, F., Muscat, A., and Plank, B., "Automatic Description Generation from Images: A Survey of Models, Datasets, and Evaluation Measures.", *Journal of Artificial Intelligence Research*, vol. 55, 409–442 2016

Cabral, R. and Furukawa, Y., "Piecewise planar and compact floorplan reconstruction from images", in "Conference on Computer Vision and Pattern Recognition", pp. 628–635, IEEE 2014

Chatterjee, M. and Schwing, A. G., "Diverse and coherent paragraph generation from images", in "European Conference on Computer Vision", 2018

Chelani, K., Sidhartha, C., and Govindu, V. M., "Towards Automated Floorplan Generation", 2018

Chen, S., Li, M., Ren, K., and Qiao, C., "Crowd map: Accurate reconstruction of indoor floor plans from crowdsourced sensor-rich videos", in "International conference on distributed computing systems", pp. 1–10, IEEE 2015a

Chen, X., Fang, H., Lin, T.-Y., Vedantam, R., Gupta, S., Dollár, P., and Zitnick, C. L., "Microsoft COCO captions: Data collection and evaluation server", *arXiv preprint arXiv:1504.00325* 2015b

Chhabra, A. K., "Graphic symbol recognition: An overview", in "International Workshop on Graphics Recognition", pp. 68–79 1997

Cho, K., Van Merriënboer, B., Bahdanau, D., and Bengio, Y., "On the properties of neural machine translation: Encoder-decoder approaches", *arXiv preprint arXiv:1409.1259* 2014

D. Sharma, C. Chattopadhyay and G. Harit, "A Unified Framework for Semantic Matching of Architectural Floorplans ", in "International Conference on Pattern Recognition", 2016

Dasgupta, S., Fang, K., Chen, K., and Savarese, S., "Delay: Robust spatial layout estimation for cluttered indoor scenes", in "Conference on Computer Vision and Pattern Recognition", pp. 616–624 2016

de las Heras, L.-P., Fernández, D., Valveny, E., Lladós, J., and Sánchez, G., "Unsupervised wall detector in architectural floor plans", in "International Conference on Document Analysis and Recognition", pp. 1245–1249, IEEE 2013

de las Heras, L.-P., Ahmed, S., Liwicki, M., Valveny, E., and Sánchez, G., "Statistical segmentation and structural recognition for floor plan interpretation", *International Journal on Document Analysis and Recognition*, vol. 17, no. 3, 221–237 2014

de las Heras, L.-P., Terrades, O. R., Robles, S., and Sánchez, G., "CVC-FP and SGT: A new database for structural floor plan analysis and its groundtruthing tool", *International Journal on Document Analysis and Recognition*, vol. 18, no. 1, 15–30 2015

Delalandre, M., Pridmore, T., Valveny, E., Locteau, H., and Trupin, E., "Building synthetic graphical documents for performance evaluation", in "International Workshop on Graphics Recognition", pp. 288–298, Springer 2007

Delalandre, M., Valveny, E., Pridmore, T., and Karatzas, D., "Generation of synthetic documents for performance evaluation of symbol recognition & spotting systems", *International Journal on Document Analysis and Recognition*, vol. 13, no. 3, 187–207 2010

Denkowski, M. and Lavie, A., "Meteor 1.3: Automatic metric for reliable optimization and evaluation of machine translation systems", in "Workshop on Statistical Machine Translation", 2011

Dodge, S., Xu, J., and Stenger, B., "Parsing floor plan images", in "International Conference on Machine Vision Applications", pp. 358–361, IEEE 2017

Doermann, D., Tombre, K. *et al.*, *Handbook of document image processing and recognition*, Springer 2014

Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., and Darrell, T., "Long-term recurrent convolutional networks for visual recognition and description", in "Conference on Computer Vision and Pattern Recognition", pp. 2625–2634 2015

Dosch, P., Tombre, K., Ah-Soon, C., and Masini, G., "A complete system for the analysis of architectural drawings", *International Journal on Document Analysis and Recognition*, vol. 3, no. 2, 102–116 2000

Dutta, A. and Zisserman, A., "The VIA Annotation Software for Images, Audio and Video", in "International Conference on Multimedia", ACM 2019

Dutta, A., Lladós, J., and Pal, U., "A symbol spotting approach in graphical documents by hashing serialized graphs", *Pattern Recognition*, vol. 46, no. 3, 752–768 2013

Elliott, D. and Keller, F., "Image description using visual dependency representations", in "Empirical Methods in Natural Language Processing", pp. 1292–1302 2013

Elliott, D. and Keller, F., "Comparing automatic evaluation measures for image description", in "Association for Computational Linguistics", 2014

Farhadi, A., Hejrati, M., Sadeghi, M. A., Young, P., Rashtchian, C., Hockenmaier, J., and Forsyth, D., "Every picture tells a story: Generating sentences from images", in "European Conference on Computer Vision", 2010

Fernandez-Labrador, C., Perez-Yus, A., Lopez-Nicolas, G., and Guerrero, J. J., "Layouts from panoramic images with geometry and deep learning", *IEEE Robotics and Automation Letters*, vol. 3, no. 4, 3153–3160 2018

F.J. Damerau, "A technique for computer detection and correction of spelling errors", *Communications of the ACM*, vol. 7, no. 3, 171–176 1964

Fornés, A., Lladós, J., Sánchez, G., and Karatzas, D., "Rotation invariant hand-drawn symbol recognition based on a dynamic time warping model", *International Journal on Document Analysis and Recognition*, vol. 13, no. 3, 229–241 2010

Freeman, H., "Computer processing of line-drawing images", *ACM Computing Surveys*, vol. 6, no. 1, 57–97 1974

Furlan, A., Miller, S. D., Sorrenti, D. G., Li, F.-F., and Savarese, S., "Free your Camera: 3D Indoor Scene Understanding from Arbitrary Camera Motion.", in "British Machine Vision Conference", 2013

Girshick, R., "Fast R-CNN", in "International Conference on Computer Vision", 2015

Google Developers, "ARcore, fundamental concepts", `https://developers.google.com/ar/discover/concepts` 2020a

Google Developers, "Google Measure App", `https://developer.android.com/reference/kotlin/android/icu/util/Measure?hl=en` 2020b

Goyal, S., Chattopadhyay, C., and Bhatnagar, G., "ASYSST: A Framework for Synopsis Synthesis Empowering Visually Impaired", in "Workshop on Multimedia for Accessible Human Computer Interface", 2018a

Goyal, S., Chattopadhyay, C., and Bhatnagar, G., "Plan2Text: A framework for describing building floor plan images from first person perspective", in "Colloquium on Signal Processing and its Applications", 2018b

Goyal, S., Bhavsar, S., Patel, S., Chattopadhyay, C., and Bhatnagar, G., "SUGAMAN: describing floor plans for visually impaired by annotation learning and proximity-based grammar", *Image Processing*, vol. 13, no. 13, 2623–2635 2019a

Goyal, S., Mistry, V., Chattopadhyay, C., and Bhatnagar, G., "BRIDGE: Building plan Repository for Image Description Generation, and Evaluation", in "International Conference on Document Analysis and Recognition", 2019b

Goyal, S., Mistry, V., Chattopadhyay, C., and Bhatnagar, G., "BRIDGE: Building Plan Repository for Image Description Generation, and Evaluation", in "International Conference on Document Analysis and Recognition", pp. 1071–1076, IEEE 2019c

He, K., Zhang, X., Ren, S., and Sun, J., "Spatial pyramid pooling in deep convolutional networks for visual recognition", *Transaction on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, 1904–1916 2015

He, K., Gkioxari, G., Dollár, P., and Girshick, R., "Mask r-cnn", in "International Conference on Computer Vision", 2017

Hilaire, X. and Tombre, K., "Robust and accurate vectorization of line drawings", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, 890–904 2006

Hochreiter, S. and Schmidhuber, J., "Long short-term memory", *Neural computation*, vol. 9, no. 8, 1735–1780 1997

Hodosh, M., Young, P., and Hockenmaier, J., "Framing image description as a ranking task: Data, models and evaluation metrics", *Journal of Artificial Intelligence Research*, vol. 47, 853–899 2013

Hsiao, C.-W., Sun, C., Sun, M., and Chen, H.-T., "Flat2Layout: Flat Representation for Estimating Layout of General Room Types", *arXiv preprint arXiv:1905.12571* 2019

Hu, M. K., "Visual pattern recognition by moment invariants, computer methods in image analysis", *Transations on Information Theory*, vol. 8 1962

Jain, A. K. and Yu, B., "Document representation and its application to page decomposition", *Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, 294–308 1998

Johnson, J., Karpathy, A., and Fei-Fei, L., "Densecap: Fully convolutional localization networks for dense captioning", in "Conference on Computer Vision and Pattern Recognition", 2016

K. Tombre and S. Tabbone and L. Pélissier, B. Lamiroy and P. Dosch, "Text/graphics separation revisited", in "International Workshop on Document Analysis Systems", 2002

Kafle, K., Price, B., Cohen, S., and Kanan, C., "DVQA: Understanding data visualizations via question answering", in "Conference on computer vision and pattern recognition", pp. 5648–5656, IEEE 2018

Kahou, S. E., Michalski, V., Atkinson, A., Kádár, Á., Trischler, A., and Bengio, Y., "Figureqa: An annotated figure dataset for visual reasoning", *arXiv preprint arXiv:1710.07300* 2017

Kalervo, A., Ylioinas, J., Häikiö, M., Karhu, A., and Kannala, J., "Cubicasa5k: A dataset and an improved multi-task model for floorplan image analysis", in "Scandinavian Conference on Image Analysis", pp. 28–40, Springer 2019

Karpathy, A. and Fei Fei, L., "Deep visual-semantic alignments for generating image descriptions", in "Conference on Computer Vision and Pattern Recognition", 2015

Kasturi, R., O'Gorman, L., and Govindaraju, V., "Document image analysis: A primer", *Sadhana*, vol. 27, no. 1, 3–22 Feb 2002

Kembhavi, A., Seo, M., Schwenk, D., Choi, J., Farhadi, A., and Hajishirzi, H., "Are you smarter than a sixth grader? textbook question answering for multimodal machine comprehension", in "Conference on Computer Vision and Pattern recognition", pp. 4999–5007, IEEE 2017

Kilmer, W. O. and Kilmer, R., *Construction drawings and details for interiors: Basic skills*, John Wiley & Sons 2009

Kiros, R., Salakhutdinov, R., and Zemel, R. S., "Unifying visual-semantic embeddings with multimodal neural language models", *arXiv preprint arXiv:1411.2539* 2014

Krause, J., Johnson, J., Krishna, R., and Fei-Fei, L., "A hierarchical approach for generating descriptive image paragraphs", in "Conference on Computer Vision and Pattern Recognition", 2017

Krishna, R., Zhu, Y., Groth, O., Johnson, J., Hata, K., Kravitz, J., Chen, S., Kalantidis, Y., Li, L.-J., Shamma, D. A. *et al.*, "Visual genome: Connecting language and vision using crowdsourced dense image annotations", *International Journal of Computer Vision*, vol. 123, no. 1, 32–73 2017

Kulkarni, G., Premraj, V., Dhar, S., Li, S., Choi, Y., Berg, A. C., and Berg, T. L., "Baby talk: Understanding and generating image descriptions", in "Conference on Computer Vision and Pattern Recognition", 2011

Kulkarni, G., Premraj, V., Ordonez, V., Dhar, S., Li, S., Choi, Y., Berg, A. C., and Berg, T. L., "Babytalk: Understanding and generating simple image descriptions", *Transaction on Pattern Analysis and Machine Intelligence*, vol. 35, no. 12, 2891–2903 2013

Kuznetsova, P., Ordonez, V., Berg, A. C., Berg, T. L., and Choi, Y., "Collective generation of natural image descriptions", in "Association for Computational Linguistics", 2012

L.A. Fletcher and R. Kasturi, "A robust algorithm for text string separation from mixed text/graphics images", *Transaction on Pattern Recognition and Machine Analysis*, vol. 10, no. 6, 910–918 1988

Lai, C. and Kasturi, R., "Detection of dashed lines in engineering drawings and maps", in "International Conference on Document Analysis and Recognition", 1991

Le Bodic, P., Héroux, P., Adam, S., and Lecourtier, Y., "An integer linear program for substitution-tolerant subgraph isomorphism and its use for symbol spotting in technical drawings", *Pattern Recognition*, vol. 45, no. 12, 4214–4224 2012

Li, S., Kulkarni, G., Berg, T. L., Berg, A. C., and Choi, Y., "Composing simple image descriptions using web-scale n-grams", in "Conference on Computational Natural Language Learning", 2011

Lin, C., Li, C., Furukawa, Y., and Wang, W., "Floorplan priors for joint camera pose and room layout estimation", *arXiv preprint arXiv:1812.06677* 2018

Lin, C.-Y., "Rouge: A package for automatic evaluation of summaries", in "Association for Computational Linguistics", 2004

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L., "Microsoft coco: Common objects in context", in "European Conference on Computer Vision", 2014

Liu, C., Schwing, A. G., Kundu, K., Urtasun, R., and Fidler, S., "Rent3d: Floor-plan priors for monocular layout estimation", in "Conference on Computer Vision and Pattern Recognition", pp. 3413–3421 2015

Liu, L. and Zlatanova, S., "A "door-to-door" path-finding approach for indoor navigation", *GeoInformation for Disaster Management* 2011

Liu, Y., Fu, J., Mei, T., and Chen, C. W., "Let your photos talk: Generating narrative paragraph for photo stream via bidirectional attention recurrent neural networks", in "Association for the Advancement of Artificial Intelligence", 2017

Luong, M.-T., Pham, H., and Manning, C. D., "Effective approaches to attention-based neural machine translation", *arXiv preprint arXiv:1508.04025* 2015

Macé, S., Locteau, H., Valveny, E., and Tabbone, S., "A system to detect rooms in architectural floor plan images", in "International Workshop on Document Analysis Systems", pp. 167–174, ACM 2010

Madugalla, A., Marriott, K., Marinai, S., Capobianco, S., and Goncu, C., "Creating Accessible Online Floor Plans for Visually Impaired Readers", *Transactions on Accessible Computing*, vol. 13, no. 4, 1–37 2020

Mao, Y., Zhou, C., Wang, X., and Li, R., "Show and Tell More: Topic-Oriented Multi-Sentence Image Captioning.", in "International Joint Conferences on Artificial Intelligence", 2018

Marcus, M. P., Marcinkiewicz, M. A., and Santorini, B., "Building a large annotated corpus of English: The Penn Treebank", *Computational linguistics*, vol. 19, no. 2, 313–330 1993

Mello, C. A., Costa, D. C., and dos Santos, T., "Automatic image segmentation of old topographic maps and floor plans", in "International Conference on Systems, Man and Cybernetics", 2012

Methani, N., Ganguly, P., Khapra, M. M., and Kumar, P., "Plotqa: Reasoning over scientific plots", in "Winter Conference on Applications of Computer Vision", pp. 1527–1536, IEEE 2020

Mondal, A. and Jawahar, C., "Textual Description for Mathematical Equations", in "International Conference on Document Analysis and Recognition", pp. 1300–1307, IEEE 2019

Murali, S., Speciale, P., Oswald, M. R., and Pollefeys, M., "Indoor scan2bim: Building information models of house interiors", in "International Conference on Intelligent Robots and Systems", pp. 6126–6133, IEEE 2017

Nallapati, R., Zhou, B., Gulcehre, C., Xiang, B. *et al.*, "Abstractive text summarization using sequence-to-sequence rnns and beyond", *arXiv preprint arXiv:1602.06023* 2016

Nayef, N. and Breuel, T. M., "Statistical grouping for segmenting symbols parts from line drawings, with application to symbol spotting", in "International Conference on Document Analysis and Recognition", pp. 364–368, IEEE 2011

Occipital, "TapeMeasure: The fastest way to measure (iOS application)", `https://tapmeasure.io/` 2020

O'Gorman, L. and Kasturi, R., *Document image analysis*, vol. 39, IEEE Computer Society Press Los Alamitos 1995

Ojala, T., Pietikainen, M., and Maenpaa, T., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *Transaction on Pattern Recognition and Machine Analysis*, vol. 24, no. 7, 971–987 2002

Okorn, B., Xiong, X., Akinci, B., and Huber, D., "Toward automated modeling of floor plans", in

"Symposium on 3D data processing, visualization and transmission", vol. 2 2010

Ordonez, V., Kulkarni, G., and Berg, T. L., "Im2text: Describing images using 1 million captioned photographs", in "Conference on Neural Information Processing Systems", 2011

Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J., "BLEU: a method for automatic evaluation of machine translation", in "Association for Computational Linguistics", 2002

Park, C. C. and Kim, G., "Expressing an image stream with a sequence of natural sentences", in "Neural Information Processing Systems", 2015

Phalak, A., Badrinarayanan, V., and Rabinovich, A., "Scan2Plan: Efficient Floorplan Generation from 3D Scans of Indoor Scenes", *arXiv preprint arXiv:2003.07356* 2020

Qureshi, R. J., Ramel, J.-Y., Barret, D., and Cardot, H., "Spotting symbols in line drawing images using graph representations", in "International Workshop on Graphics Recognition", pp. 91–103, Springer 2007

Reddy, R., Ramesh, R., Deshpande, A., and Khapra, M. M., "Figurenet: A deep learning model for question-answering on scientific plots", in "International Joint Conference on Neural Networks", pp. 1–8, IEEE 2019

Redmon, J. and Farhadi, A., "YOLO9000: better, faster, stronger", in "Conference on Computer Vision and Pattern Recognition", 2017

Redmon, J. and Farhadi, A., "Yolov3: An incremental improvement", *arXiv preprint arXiv:1804.02767* 2018

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., "You only look once: Unified, real-time object detection", in "Conference on Computer Vision and Pattern Recognition", 2016

Ren, S., He, K., Girshick, R., and Sun, J., "Faster R-CNN: Towards real-time object detection with region proposal networks", in "Conference on Neural Information Processing Systems", 2015

Rezvanifar, A., Cote, M., and Branzan Albu, A., "Symbol Spotting on Digital Architectural Floor Plans Using a Deep Learning-based Framework", in "Conference on Computer Vision and Pattern Recognition Workshops", pp. 568–569 2020

Rush, A. M., Chopra, S., and Weston, J., "A neural attention model for abstractive sentence summarization", *arXiv preprint arXiv:1509.00685* 2015

Rusiñol, M. and Lladós, J., "State-of-the-Art in Symbol Spotting", in "Symbol Spotting in Digital Libraries", pp. 15–47, Springer 2010

Rusiñol, M., Borràs, A., and Lladós, J., "Relational indexing of vectorial primitives for symbol spotting in line-drawing images", *Pattern Recognition Letters*, vol. 31, no. 3, 188–201 2010

S. Ahmed, M. Liwicki, M. Weber and A. Dengel, "Text/graphics segmentation in architectural floor plans", in "International Conference on Document Analysis and Recognition", 2011

S. Goyal, C. Chattopadhyay, and G. Bhatnagar, "Plan2Text: A Framework for Describing Building Floor Plan Images From First Person Perspective", in "Colloquium on Signal Processing and its Applications", 2018

Sabour, S., Frosst, N., and Hinton, G. E., "Dynamic routing between capsules", in "Advances in neural information processing systems", pp. 3856–3866 2017

Saha, R., Mondal, A., and Jawahar, C., "Graphical Object Detection in Document Images", in "International Conference on Document Analysis and Recognition", pp. 51–58, IEEE 2019

Satorra, A. and Bentler, P. M., "A scaled difference chi-square test statistic for moment structure analysis", *Psychometrika*, vol. 66, no. 4, 507–514 2001

Schreiber, S., Agne, S., Wolf, I., Dengel, A., and Ahmed, S., "Deepdesrt: Deep learning for detection and structure recognition of tables in document images", in "International Conference on Document Analysis and Recognition", vol. 1, pp. 1162–1167, IEEE 2017

Sensopia, " MagicPlan: Create a floor plan in just a few minutes", `https://www.magicplan.app/magicplan/` 2020

Sharma, D., Gupta, N., Chattopadhyay, C., and Mehta, S., "DANIEL: A deep architecture for automatic analysis and retrieval of building floor plans", in "International Conference on Document Analysis and Recognition", vol. 1, pp. 420–425, IEEE 2017

Sharma, N., Mandal, R., Sharma, R., Pal, U., and Blumenstein, M., "Signature and Logo Detection using Deep CNN for Document Image Retrieval", in "International Conference on Frontiers of Handwriting Recognition", 2018

Sun, C., Hsiao, C.-W., Sun, M., and Chen, H.-T., "Horizonnet: Learning room layout with 1d representation and pano stretch data augmentation", in "Conference on Computer Vision and Pattern Recognition", pp. 1047–1056 2019

Sutskever, I., Vinyals, O., and Le, Q., "Sequence to sequence learning with neural networks", *Conference on Neural Information Processing Systems* 2014

Tigora, A. *et al.*, "An overview of document image analysis systems", *EDITURA UNIVERSITARA Bucuresti*, p. 378 2013

Trier, O. D., Taxt, T., and Jain, A. K., "Data capture from maps based on gray scale topographic analysis", in "International Conference on Document Analysis and Recognition", vol. 2, pp. 923–926 1995

Turner, E. and Zakhor, A., "Floor plan generation and room labeling of indoor environments from laser range data", in "International conference on computer graphics theory and applications", pp. 1–12, IEEE 2014

Tzutalin, "Tzutalin. LabelImg. Git code (2015).", `https://github.com/tzutalin/labelImg` 2015

V. Yadav and N. Ragot, "Text extraction in document images: highlight on using corner points", in "International Workshop on Document Analysis Systems", 2016

Valveny, E. and Martí, E., "Application of deformable template matching to symbol recognition in handwritten architectural drawings", in "International Conference on Document Analysis and Recognition", pp. 483–486, IEEE 1999

Vedantam, R., Lawrence Zitnick, C., and Parikh, D., "Cider: Consensus-based image description evaluation", in "Conference on Computer Vision and Pattern Recognition", pp. 4566–4575 2015

Viola, P. and Jones, M., "Rapid object detection using a boosted cascade of simple features", in "Conference on Computer Vision and Pattern Recognition", 2001

Wang, Q. and Chan, A. B., "CNN+ CNN: convolutional decoders for image captioning", *arXiv preprint arXiv:1805.09019* 2018

Wang, Z., Luo, Y., Li, Y., Huang, Z., and Yin, H., "Look Deeper See Richer: Depth-aware Image Paragraph Captioning.", in "ACM Multimedia", 2018

Xu, J., Stenger, B., Kerola, T., and Tung, T., "Pano2CAD: Room layout from a single panorama image", in "Winter Conference on Applications of Computer Vision", pp. 354–362, IEEE 2017

Xu, M., Wei, S., and Zlatanova, S., "An indoor navigation approach considering obstacles and space subdivision of 2D plan", *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 41, 339 2016

Yao, T., Pan, Y., Li, Y., Qiu, Z., and Mei, T., "Boosting image captioning with attributes", in "International Conference on Computer Vision", 2017

Yi, X., Gao, L., Liao, Y., Zhang, X., Liu, R., and Jiang, Z., "CNN based page object detection in document images", in "International Conference on Document Analysis and Recognition", 2017

Z. Lu, "Detection of text regions from digital engineering drawings", *Transaction on Pattern Recognition and Machine Analysis*, vol. 20, no. 4, 431–439 1998

Zeng, Z., Li, X., Yu, Y. K., and Fu, C.-W., "Deep Floor Plan Recognition Using a Multi-Task Network with Room-Boundary-Guided Attention", in "International Conference on Computer Vision", pp. 9096–9104 2019

Zhang, J., Kan, C., Schwing, A. G., and Urtasun, R., "Estimating the 3d layout of indoor scenes and its clutter from depth sensors", in "International Conference on Computer Vision", pp. 1273–1280 2013

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., and Artzi, Y., "Bertscore: Evaluating text generation with bert", *arXiv preprint arXiv:1904.09675* 2019a

Zhang, W., Zhang, W., and Gu, J., "Edge-semantic learning strategy for layout estimation in indoor environment", *IEEE Transaction on Cybernetics* 2019b

Zhang, Y., Song, S., Tan, P., and Xiao, J., "Panocontext: A whole-room 3d context model for panoramic scene understanding", in "European Conference on Computer Vision", pp. 668–686, Springer 2014

Ziran, Z. and Marinai, S., "Object Detection in Floor Plan Images", in "Workshop on Artificial Neural Networks in Pattern Recognition", 2018

Zlatanova, S., Liu, L., and Sithole, G., "A conceptual framework of space subdivision for indoor navigation", in "International Workshop on Indoor Spatial Awareness", pp. 37–41, ACM 2013

Zou, C., Colburn, A., Shan, Q., and Hoiem, D., "Layoutnet: Reconstructing the 3d room layout from a single rgb image", in "Conference on Computer Vision and Pattern Recognition", pp. 2051–2059 2018