Depth Estimation in 2D images

*Note: Sub-titles are not captured in Xplore and should not be used

Nikhil Chapre
Vellore Institute of Technology
Chennai, India
nikhilc2209@gmail.com

Shaunak Deshpande
Vellore Institute of Technology
Chennai, India
shaunak.deshpande4@gmail.com

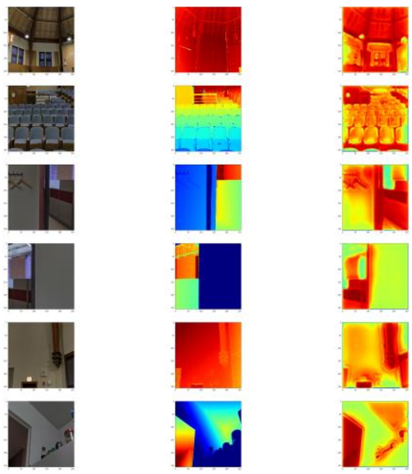
***Abstract*— This study is based on a Deep Learning approach to generate depth maps based on input 2d images. We train a CNN using the DIODE depth dataset and use it to generate depth maps and compare them with existing depth maps from the training set. This study uses dataset having both indoor and outdoor images from various scenes having varying depth.**

Keywords—deep learning, depth, CNN, DIODE dataset, insert

I. Introduction (*Heading 1*)

Depth estimation in monocular images often poses to be a challenging task as without any perception of depth it is difficult to visualize distance of objects from the screen. It is easier in stereoscopic images because of multiple images from several camera angles. Since depth estimation in general has its various applications ranging from scene reconstruction for 3D printing or just for the sake of visualizations.To tackle this problem we try to use Supervised Learning to train models which can accurately identify the distance of objects from the screen(camera) and form depth maps for each pixel for easier visualization. Using convolutional neural networks we can form depth maps to assign different colours based on a colormap gradient scale. For this purpose we will make use of the DIODE indoor dataset which contains high quality indoor images with their respective depth maps and masks.

Using a supervised learning method will help us in understanding the difference between a actual depth and the depth map predicted by the model.



II. Literature Review

A. Object Detection of Remote Sensing Airport Image Based on Improved Faster R-CNN: HAN Yongsai1, MA Shiping2, ZHANG Fei, LI Chenghao. The model used here was an efficient version of Region based Convolutional Neural Network. The use case was remote sensing of airport images and detecting objects in the same. A deeper basic ResNet was used, and add a new fully connected layer to the end detection network and combine the SoftMax classifier and 4 logistic regression classifiers for object detection according to the inter-class correlation of the object.

*Satellite images of Airports were used. The faster RCNN model uses Region Proposal Net (RPN) which results in more accuracy, efficiency. This framework identified objects such as airports, civil aviation aircraft, fighter jets, transport planes, helicopters, tanks, and bridge objects. Each proposed ROI is convoluted with 3*3*n filters, pooling operation used on feature map obtained from convolution of last layer.*

B. *Real-Time Depth Estimation from 2D Images: Jack Zhu, Ralph Ma, They proposed model here is Transfer learning on real time images captured by camera, study the depth maps and estimate depths using heatmap on the NYU-dataset. They compared the transfer learning model on a pretrained and untrained network, compared performance differences. The lossy function used here is RMSE. Different pretrained model were used Mean, Make3d, Course Network, Course + Fine Network, Hierarchical VGG. Hierarchical VGG produced the best efficiency and the lowest RMSE values.*

C. *Fast Automatic Vehicle Annotation for Urban Traffic Surveillance: Detection and Annotation for Vehicles (DAVE) model was suggested which simultaneously runs 2 CNNs. 1st CNN is responsible for Vehicle detection. 2nd CNN is responsible for Annotation. A shallow Fully Convolutional Fast Vehicle Proposal Network was suggested for extracting all vehicles pose, color, type, simultaneously. These two nets are jointly optimized so that abundant latent knowledge learned from the deep empirical ALN can be exploited to guide training the much simpler FVPN. The proposed DAVE framework outperforms RCNN, Fast RCNN, Faster RCNN.*

D. *Depth Prediction from 2D Images: A Taxonomy and an Evaluation Study. They performed an Evaluation study by training the data on both supervised and unsupervised models. They compared multi view and Monocular models. The non-deep learning methods use stereo matching and have some disadvantages such as needing train data while runtime. The monocular is divided into Bootstrap and Iterative networks whose output is combined into a Refined network. The monocular 2D depth maps generated are refined and used as outputs.*

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. PROPOSED WORK

IV. EXPERIMENTAL SETUP

To run our Keras models on DIODE dataset, we used Nvidia CUDA Development toolkit 11.2 on Laptop RTX 3060 which has around 3840 CUDA cores. The CPU used is AMD Ryzen 5800H which features eight SMT-enabled Zen 3 cores running at 3.2 GHz (base clock speed) to 4.4 GHz (Boost). Tegra TK1 Development kit. The TK1 provides an NVIDIA Kepler GPU with 192 CUDA cores, and a QuadCode ARM Cortex CPU. The board has 2GB RAM, split between GPU and CPU. We run Ubuntu 14.04 L4T, CUDA and CUDNN versions 6.5. To avoid hitting memory limitations, we added a 4GB swapfile on a high speed SD card.

V. KEEP YOUR TEXT AND GRAPHIC FILES SEPARATE UNTIL AFTER THE TEXT HAS BEEN FORMATTED AND STYLED. DO NOT USE HARD TABS, AND LIMIT USE OF HARD RETURNS TO ONLY ONE RETURN AT THE END OF A PARAGRAPH. DO NOT ADD ANY KIND OF PAGINATION ANYWHERE IN THE PAPER. DO NOT NUMBER TEXT HEADS-THE TEMPLATE WILL DO THAT FOR YOU.

A. *Abbreviations and Acronyms*

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. *Units*

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as "3.5-inch disk drive".
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: "Wb/m2" or "webers per square meter", not "webers/m2". Spell out units when they appear in text: ". . . a few henries", not ". . . a few H".
- Use a zero before decimal points: "0.25", not ".25". Use "cm3", not "cc". (*bullet list*)

C. *Equations*

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \tag{1}$$

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(1)", not "Eq. (1)" or "equation (1)", except at the beginning of a sentence: "Equation (1) is . . ."

Identify applicable funding agency here. If none, delete this text box.

D. Some Common Mistakes

- The word "data" is plural, not singular.
- The subscript for the permeability of vacuum μ_0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter "o".
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an "inset", not an "insert". The word alternatively is preferred to the word "alternately" (unless you really mean something that alternates).
- Do not use the word "essentially" to mean "approximately" or "effectively".
- In your paper title, if the words "that uses" can accurately replace the word "using", capitalize the "u"; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones "affect" and "effect", "complement" and "compliment", "discreet" and "discrete", "principal" and "principle".
- Do not confuse "imply" and "infer".
- The prefix "non" is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the "et" in the Latin abbreviation "et al.".
- The abbreviation "i.e." means "that is", and the abbreviation "e.g." means "for example".

An excellent style manual for science writers is [7].

VI. Using the Template

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

A. Authors and Affiliations

The template is designed for, but not limited to, six authors. A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

1) For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

2) For papers with less than six authors: To change the default, adjust the template as follows.

a) Selection: Highlight all author and affiliation lines.

b) Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

c) Deletion: Delete the author and affiliation lines for the extra authors.

B. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is "Heading 5". Use "figure caption" for your Figure captions, and "table head" for your table title. Run-in heads, such as "Abstract", will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named "Heading 1", "Heading 2", "Heading 3", and "Heading 4" are prescribed.

C. Figures and Tables

a) Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation "Fig. 1", even at the beginning of a sentence.

TABLE I. TABLE TYPE STYLES

Table Head	**Table Column Head**		
	Table column subhead	***Subhead***	***Subhead***
copy	More table copy[a]		

[a.] Sample of a Table footnote. (*Table footnote*)

Fig. 1. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity "Magnetization", or "Magnetization, M", not just "M". If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write "Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

ACKNOWLEDGMENT *(Heading 5)*

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi TIFF or EPS file, with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord "Format" pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.