Real Time Hand Gesture Recognition in Industry

With the 4th industrial revolution and the increased use of cobots in the industries comes many opportunities for new generation control panels. In this article, we proposed to develop a deep learning model to recognize in real time 10 different gestures that can be used to interact with a cobot. We proposed a new dataset containing gestures that can be used in industrial context. The videos were taken from a computer webcam and then processed to remove the noise created by the background by isolating the movement of the gray scale images. We proposed to extract the spatio-temporal features by the combination of 3D convolution and LSTM layers. We also proposed a real time method to recognize our gestures. Our experimental results show for 8 out of 10 gestures, a recognition rate of more than 90\%. Furthermore, an interface was created to test our method in real time and to add a new class of gestures to be recognized by our model.

CCS CONCEPTS • Computing methodologies ~Artificial intelligence ~Computer vision ~Computer vision tasks ~Activity recognition and understanding

(Il y a des explications sur le ccs concepts dans le template, mais je t’avoue ne pas comprendre grand-chose. J’ai quand même copié les key words ici et j’ai mis le code dans les commentaires du doc)

1. Introduction

In more recent years, the world of industry has witnessed a major change, referred to as "Industry 4.0". In fact, robots are now used in crowded environments and in shared areas with operators, the thing that demands a focus on the development of communication techniques with robots like gesture recognition for human robot interaction.

There are different levels of human robot interaction according to its degree: coexistence, synchronization, cooperation, and collaboration. Collaboration is the highest level of interaction among these levels where robot and operator share the same space and may work simultaneously on the same component. In this context, we need a specific algorithm able to recognize gestures needed to control a cobot without a control panel.

Existing human gesture datasets mostly contain gestures for sign language recognition or virtual reality context with a clear lack of industrial hand gesture recognition datasets and methods.

This paper proposed to bring a solution in this context. Indeed, we propose an interface that we used to acquire a new dataset with specific gestures that can be used in industrial context. The collected dataset is then used within a convolutional approach to extract spatio-temporal features in order to classify gestures. Moreover, we proposed a real time approach to recognize gestures using a friendly interface. The interface mainly shows a feed-back of the camera with the predictions from the model. There is also the possibility to create new videos and add new gestures in order to enrich the dataset with desired gestures.

The rest of the paper is organized as follows: Section II

examines state of the art methods for gesture recognition

And used datasets. Section III introduce

the proposed dataset that we created in industrial context. In Section IV, we will explain

Proposed approach for gesture recognition. Section V

presents the experimental results. Section VI explains how to recognize gestures in real time. Finally, the last section concludes the paper with future research directions.

1. State-of-the-art

Il faut changer tous les « cites » en références word…

Many different techniques were developed in the past years for hand gesture recognition in general. One of the first was an approach that separates the fingers from the palm \cite{chen2014real}. The position of the hand on the static image was then deduced depending on the number of fingers found by the algorithm and their relative position.

A much more developed method is to apply a skeletonization method \cite{ogniewicz1992voronoi}, \cite{ionescu2005dynamic}. The idea is to reduce an object to lines that are only one pixel wide. The images are therefore easier to analyse by the model. This method is applied to the videos before the training and the tests. An algorithm to create the skeleton of a shape need a Voronoi tessellation \cite{beristain2009skeleton} but it's a heavy algorithm that slows down the entire process. A solution would be to use the champfer distance transform \cite{butt1998optimum} to reduce the computational cost.

N. Luthfil Hakim, T. K. Shih et al \cite{hakim2019dynamic} created a model of gesture recognition for a television application. They used an RGB and a depth camera to get more valuable information. Their model is composed by multiple blocs of 3D convolution layers combined with dense and max pooling layers. The last part of their model is the Long Short-Term Memory layers \cite{hochreiter1997long} that finds features that are time related which are crucial when dealing with videos.

Another basic method is the haar-like features \cite{chen2007real} that are used for face or eyes detections. This method is called a hand crafted feature extraction method. If the color changes are not sharp enough, this method will show no results.

The LSTM layer is an advanced type of RNN \cite{giles1994dynamic}. Enhanced methods of Recurrent Neural Network that could be used for gesture recognition were developed : Differencial RNN \cite{veeriah2015differential}, Hierarchical RNN \cite{du2015hierarchical}, Bidirectionnal RNN \cite{pigou2018beyond}.

Some articles\cite{hakim2019dynamic}, \cite{ren2011robust} use the kinect camera to get the depth information in addition to the 2D images.

Transformers \cite{vaswani2017attention}, \cite{de2020sign} are a type of layers that is designed to handle sequential data. Unlike LSTMs, they use a mechanism to give different weighs to the input data.

1. Proposed dataset
   1. Dataset creation protocol

In order to recognize gestures in industrial context and for cobot controlling with gestures, we decided to create a new dataset. In order to collect this dataset, we created a program with a friendly interface that enables us to easily register videos with hand gestures.

The program is made to create a lot of data in the smallest amount of time. The user can switch between different labels by pressing a key and can start a new video with another key. See details on Figure \ref{fig:new\_video\_program}. There is a delay of one second before the video is taken to give the time to the user to place himself. The computer's camera has a frame rate of 25 FPS so each video lasts 2.4 seconds and contains 61 frames. With a good rhythm, a new video can be made every 5 seconds.

A dataset needs recordings of different people otherwise the model would only recognise the gesture of the person that did all of the data set videos.

The smallest data set needs at least 12 different people and each one does each gesture twice.

This dataset has 10 different gestures which make a minimum of 240 videos for the training set.

The result is a training set of almost 1 500 videos.

The data set is not split into a training and a test group. The data set in used in full for the training. Before the subjects recorded their movement, they were told to do the gesture the way that they wanted, with no other specification then the name of the movement and the time that they had to do it. However when we reviewed the videos, it appeared that they were all very similar regarding the speed and position of the gesture in the image. We decided to keep all of these videos in the training set and to create ourselves the test set to have more challenging videos. The gestures are done slower or faster than the training set and in various places on the screen.

* 1. Data processing

The videos from the data set have 61 frames with a dimension of 640 pixels wide by 480 pixels height and 3 layers of depths for the red, green and blue channels.

First step, conversion from RGB to gray image: Each pixel of the new gray image is calculated with the colors channels of source image with this formula:

\(Gray \leftarrow 0.299 \* Red + 0.587 \* Green + 0.114 \* Blue.\)

The second step is to isolate the movement in the images. To do that, each image is compared with the next one. The movement of the last frame can't be isolated because there are no more frame to do the comparison. That make 60 frames of movement.

The pros of keeping only the movement is to remove the background. Of course, to make the training data, the background need to be still, otherwise it will be considered as part of the gesture. This technique still reduces the noise that can be produced by the background.

The model doesn't get 60 frames for each video because it would make an over complicated model. The third step is to isolate a sample of the frames. The choice was made to take only 10 fps, that make a sample of 24 frames.

The fourth step is to resize and normalize the videos. The normalization take all the pixels and map them from 0-255 to 0-1.

After all the processing, the videos have 24 frames with a dimension of 120 pixels wide by 90 pixels height. All the pixels have a value between 0 and 1. See figure \ref{fig:normalization} for comparison.

The data set is available on the github page : .The data is already processed like previously explained.

1. Proposed approach

In order to recognize gestures on our collected dataset, we first proposed to test the model based on the work of N. L. Hakim, T. K. Shih, et al\cite{hakim2019dynamic}. In this article, data are collected from an RGB and a depth camera. The images from the two cameras are processed separately and combined after the LSTM units \cite{ullah2017action}. The image processing is made with multiple 3D convolution layers.

The authors proposed to use a data set with thousands of videos and got a result of 91\%.

Our setup does not include a depth camera. The RGB images are therefore processed on their own.

As expected, testing this approach on our data was not concluding. The problem was because of the small data set we are using and the large amount of parameters proposed in this model. Thus, we simplified it to create another model more efficient in this context.

We trained 15 different models, each one had a small difference compare to the last one in order to understand how each parameter affects the model.

Our final model that we selected is composed of 3 blocs: The first two are an arrangement of Convolution 3d, dense, dropout, maxpooling 3d and batch normalization. The last one has two ConvLSTM2D followed by a long-short term memory and a dense layer to have as many outputs as we have classes.

* 1. 3D Convolution

In the convolution layers \cite{ji20123d}, \cite{tran2015learning}, new channels are created with different kernels. A kernel is a squared matrix, whose size is generally 3x3 or 5x5.

To analyse videos, we chose to use 3D convolution layers, with a size of 3x3x3 or 5x5x5.

The goal of this layers is to find the features of the images in the video.

* 1. Dense

The dense layer is also called fully-connected layer. Each node of a layer is connected to all the nodes of the next layer. In our case a node is an entire image. Each channel of the last dimension is also an image. For example, the dimensions of the first dense layer are (24x90x120x256) which stands for 24 moments in time, 90 pixels height by 120 pixels wide and 256 images. The different images for one moment in time are created by the convolution layer.

The connection is defined by a weight with a value between minus one and one. In this architecture, the dense layer is placed after a convolution one. We can compare the dense layer to a filter. The meaningful channels created in the 3D convolution layer have a weight with a high value, the useless ones have a weight close to zero.

* 1. Dropout

When this layer is placed after a dense layer, it removes a certain percentage of the weights by giving them new random values. This layer is added in order to reduce the risk of overfitting.

* 1. Max pooling 3D

This layer reduces the size of the data set by keeping the max value of a section of an image (or multiple images when it's 3D). A 3D Max pooling layer with a matrix (3x3x3): The dimension go from (24,90,120,256) to (24,30,40,85) because the three last dimensions are divided by 3. This layer simplifies the data by keeping only the meaningful pixels.

* 1. Batch normalization

The layer transforms inputs so that they are standardized. This means that they have a mean value of zero and a standard deviation of one.

* 1. LSTM

Long Short-Term Memory \cite{hochreiter1997long} layers are a type of recurrent neural network capable of finding time dependent features.

This layer (and the ConvLSTM) is very important in this application because there are needed to differentiate "Swipe to the left" and "Swipe to the right" for example.

* 1. ConvLSTM 2D

Like the classic LSTM, the convLSTM is capable of learning order dependence in sequence but this layer is specialized for the videos analyses because it's the combination of LSTM and convolution layer.

In this architecture, two ConvLSTM 2D are stacked.

The first ConvLSTM layer provides an output for each input while the second layer gives only one output for a sequence of input.

* 1. Reshape

The layer doesn't learn anything in the training process. It only changes the dimension of the data to pass them to the LSTM layer which needs specific input.

1. Experiments results

The videos are stored in a raw way, they need to be processed when they are loaded to train the model. The data set of 1 500 videos is too heavy to be loaded in a standard computer's RAM, so the videos have to be load by package.

The label of the videos are saved in a list which is randomly mixed. The videos can then be loaded in the order of the list. It's important that the model is trained on different gestures every time.

To train the model faster, the data processing and training are on two different threads. The process thread loads two packs of 40 videos each and waits for the training part to be done with the previous pack. Then the loaded pack is sent to the training section. The videos are loaded by small packs of 40, when the model finishes to train on them they are unloaded to avoid overloading the computer's memory.

For each pack of videos, there are multiple batch to train on: we chose a batch size of 20 videos.

UNE TABLE ICI

The first two gestures are often confused with one another. This is probably due to the proximity of theses gesture. We can see that closing a hand and opening a hand is done at the same places on the videos. Furthermore, the difference between closing and opening a hand is much more subtle than the difference between swipe left and swipe right.

For all the other gestures, the model is very efficient. On the live test, we often got a precision of +90\%.

1. Real time recognition application
   1. Real time recognition method

The application works with threads, there is one for the prediction and another one for the interface. The videos are recorded at 10 frames per seconds, which makes a pack of 24 frames, and then are sent to the prediction. 24 frames is the number of images that were given per video to train the model.The frames are processed directly when they are taken, the list of processed images is sent to the prediction thread.

A CPU Intel i7-1065G7 takes 1.5s while a GPU Nvidia gtx1050 takes 0.4s to get the prediction done.

More predictions could be done in parallel. For example send a video for prediction every 12 frames. The 12 other frames would be kept from the end of the previous video. However it could create some lag.

* 1. Interface

The interface was coded using the python libraries tkinter and openCV.

The main window, that can be seen at the Figure \ref{fig:window1}, shows the live feed from the camera. It also shows the prediction that have the highest score at the bottom of the screen. Furthermore the window shows the 3 first predictions for the same video with the percentage of recognition. If the first prediction given by the model is lower than 60\%, the predictions are not displayed and a message is written on the screen explaining that the model did not recognise the gesture. \\

The last information that is shown is the number of frames that are already captured in the video. This allows the person in front of the camera to perform the full gesture on one video. As soon as the video is fully captured, another video is being captured. \\

The main window also shows two buttons. The first one freezes the prediction. In other words, the videos are no longer captured and the prediction stays on the window. The second button opens a second window. While the second window is active, the first is completely freezed to avoid wasting the resources of the computer.

The second window, that can be seen at the Figure \ref{fig:window2}, displays an example of each gesture that the model is able to recognise with its name. It also allows the user to add a new gesture. The first step is to create a new label or title for the gesture and add it to the list through the text field on the window. If the label already exists on the list, a message is displayed on the screen. \\

On this window, there are two buttons. The first one opens program to record a video and add it to the training set. This program is explained at \nameref{sec:methods}: A. Data set. The second button retrains the model. This operation takes a lot of time. Therefore, a safety needs to be installed to avoid miss-clicks.

1. Conclusions and future works

In this paper, we proposed a dataset collected within a new friendly used interface in order to acquire hand gestures. The collected data aims to be used for hand gesture recognition for controlling cobots in industrial context. We also proposed a well adapted architecture of a deep learning approach that is able to recognize in real time such gestures with efficiency.

The result we obtained allows us to obtain an overall accuracy of 84\% and 90\% for 8 gestures out of 10.

In future works, we intend to collect more data in the industrial context using our developed interface. We also have to focus on resolving the problem of gesture confusion by improving the model performances.

1. Bibliographie et reference

(il faut lire la doc c’est p-ê différent)

1. Inserting Content Elements

The next subsections provide instructions on how to insert figures, tables, and equations in your document.

* 1. Tables

Tables are “float elements” which should be inserted after their first text reference and have specific styles for identification. Do not use images to present tables, or they will be inaccessible to readers using assistive technologies.

Authors can insert tables by using the MS Word option (INSERT ->Table) and providing the required row and column size. Every table must have a caption (title) above it, which must have the **“TableCaption**” style applied. Please note that tables **should not** be supplied as image files, but if they are images they must have the “Image” style applied. As an example, Table 1 shows all the styles available in this template, to be applied to the respective element of your text.

Table 1: Styles available in the Word template

| Style Tag | Definition | Style Tag | Definition |
| --- | --- | --- | --- |
| Title\_document | main title of article | ListParagraph | list items |
| Subtitle | subtitle of article | Statements | math statements |
| Authors | author name | Extract | block quotations |
| Affiliation | author affiliation information | Algorithm Caption | caption for algorithm |
| AuthNotes | footnote to author(s) | AckHead | heading for acknowledgements |
| Abstract | abstract text | AckPara | acknowledgements text |
| CCSHead | heading for CSS Concepts | GrantSponsor | sponsor of grant |
| CCSDescription | CSS terms | GrantNumber | number for the grant |
| KeyWordHead | heading for keywords | ReferenceHead | heading for references |
| Keywords | keywords text | Bib\_entry | references |
| ORCID | author's ORCHID # | AppendixH1 | appendix heading level 1 |
| Head1 | heading level 1 | AppendixH2 | appendix heading level 2 |
| Head2 | heading level 2 | AppendixH3 | appendix heading level 3 |
| Head3 | heading level 3 | TableCaption | title of table |
| PosHeadPara | first paragraph after a heading | TableHead  TableFootnote | column head of table  footnote to table |
| Para | Subsequent paragraphs of general text | Image | figures |
| ParaContinue  DisplayFormula | flush left text after display items like math equations, lists etc.  numbered math equation | DOI | Digital object identifier |
| DisplayFormulaUnnum | unnumbered equations | Label | label |

Tables can be very difficult for people using screen reader technology to understand unless they include markup that explicitly defines the relationships between all the parts (i.e.: headers and data cells). *A key to making data tables accessible to screen reader users is to clearly identify column and row headers.* In Word, authors should identify which row or rows contain column headers. Below are the steps to do this:

1. Select that table’s row, then right-click the row and select “Table Properties”;
2. In the *Table Properties* window, click the *Row* tab and select the box that says “Repeat as header row at the top of each page.”

Or

Apply the “table head” style by highlighting the respective row and applying the “**TableHead**” style found in the “Body Element” section of the ACM Master Article Template.

* 1. Figures

Figures are “float elements” which should be inserted after their first text reference, and have specific styles for identification. Insert a figure and apply the “**Image**” paragraph style to it. For the figure caption, apply the style “**FigureCaption.**”

To accommodate readers with color vision differences, figures should still be usable when printed in grayscale. Refer to elements of the figure with non-color terms, for example “indicated as squares” instead of “indicated in blue”. Use different patterns in bar charts, different line patterns in graphs, and different shapes in plots to distinguish groups of elements and reinforce color differences.

* + 1. Half Width Figures.

Figure 1 is an example of a figure and caption spanning the half-page width (one column in a two column format) with the styles applied. If your figure contains third-party material, you must clearly identify it as such, as shown in the example below.



Figure 1: 1907 Franklin Model D roadster. Photograph by Harris & Ewing, Inc. [Public domain], via Wikimedia Commons. (https://goo.gl/VLCRBB)

* + 1. Full Width Figures.

Figure 2 is an example of a figure and caption spanning the full-page width with the styles applied. If your figure contains third-party material, you must clearly identify it as such, as shown in the examples.



Figure 2: Mockup of a bombe machine at Bletchley Part. Photograph by Sarah Hartwell. [Public domain], via Wikimedia Commons. (<https://commons.wikimedia.org/wiki/File:TuringBombeBletchleyPark.jpg>)

* + 1. Figure Descriptions.

Every figure should have a figure description unless it is purely decorative. These descriptions convey what’s in the image to someone who cannot see it. They are also used by search engine crawlers for indexing images, and when images cannot be loaded.

A figure description must be unformatted plain text less than xxx characters long. Figure descriptions should not repeat the figure caption – their purpose is to capture important information that is not already provided in the caption or the main text of the paper. For figures that convey important and complex new information, a short plain text description may not be adequate. More complex alternative descriptions can be placed in an appendix and referenced in a short figure description. For example, provide a data table capturing the information in a bar chart, or a structured list representing a graph. For additional information regarding how best to write figure descriptions and why doing this is so important, please see [https://www.acm.org/accessibility.](https://www.acm.org/accessibility)

The instructions below describe the required steps authors need to follow in order to insert descriptive text for figures (alt-txt value) in **MS Word 2019 on Windows or Word 2016 and later on Mac**:

1. Insert a picture in the document.
2. Right-click the image and select “Edit Alt Text”.
3. In the “alt text” section, provide your text description of the image.

Below are the steps to insert figure descriptions in **MS Word 2013 and 2016**:

1. Insert a picture in the document.
2. Right click on the inserted picture and select the **Format Picture** option.
3. In the settings at the right side of the window, click on the “Layout & Properties” icon (3rd option).
4. Expand **Alt Txt** option.
5. In the “Title” and “Description” text boxes, type the text you want to represent the figure, and then click “Close.”

Below are steps to insert the alt-txt value in **MS Word 2010/2011 for Windows\***:

1. Insert a picture in the document.
2. Right click on the inserted picture and select the **Format Picture** option.
3. Select the **Alt Txt** option from the left-side panel options.
4. In the “Title” and “Description” text boxes, type the text you want to represent the picture, and then click “Close.”  
   \* The Mac 2011 version 14.0.0 and later allows the option for inserting “alt-text.” In the MAC version of Word 2016, right-click on the image and select “Edit Alt Text” from the pop-up menu and then enter the description for the alt text.
   1. Quotations and Extracts

There are styles for block quotations, which should be used for quotes that are separated from in-line text. Below is an example.

“Microsoft tried to revive the idea of an assistant with Clippy, who began popping up in Microsoft Office in 1997. Its creator, Kevan Atteberry, was actually contracted by Microsoft to design Clippy, which, funnily enough, he did on a Mac … Sure, people could disable Clippy, but the fact he was on by default angered people.” [10]

* 1. Equations

There are two types of math equations: the *numbered display math equation* and the *un-numbered display math equation*. Below are examples of both.

* + 1. *DisplayFormula.*

The **DisplayFormula** style is applied in the numbered math equation. A numbered display equation always has an equation number (label) on the right.

(1)

* + 1. DisplayFormula.Unnum.

The **DisplayFormulaUnnum** style is applied only in unnumbered equations. An unnumbered display equation never contains an equation number Bertot and Grimes (2012) on the right—this element distinguishes it from the numbered equation.

Please note: the subsequent text after the **DisplayFormula** (numbered equation) or **DisplayFormulaUnnum** (unnumbered equation) must have the paragraph style **ParaContinue** applied.

* 1. Math statements

Math statements should have the “Statement” style applied.

**Theorem/Proof/Lemma.** Math statements should have the “**Statement**” style applied. This paragraph is an example of the “**Statement**” style.

* 1. Algorithms

Algorithms use the styles “AlgorithmCaption” and “Algorithm”.

ALGORITHM 1: Iterative Algorithm

current\_position center

current\_direction up

current\_position is inside circle

while current\_position is inside circle, do

neighborhood all grid hexes within two hexes from current\_position

for each hex in neighborhood, do

for each neuron in hex do

convert neuron\_orientation to vector

scale vector by neuron\_excitation

vector\_sum vector\_sum + vector

end

end

normalize vector\_sum

end

1. Citing Related Work

This section cites a variety of journal [5, 15], conference [1, 6, 8, 12, 13], and magazine [3] articles to illustrate how they appear in the references section. It also cites books [9, 10], a technical report [7], a PhD dissertation [4], an online reference [14], a software artifact [11], and a dataset [2].

ACKNOWLEDGMENTS

Acknowledgments are placed before the references. Add information about grants, awards, or other types of funding that you have received to support your research. Author can capture the **grant sponsor information**, by selecting the grant sponsor text and apply style ‘GrantSponsor’. After this, select grant no and apply ‘GrantNumber’ from style panel. Example of Grant sponsor: Competitive Research Programme and example of Grant no: CRP 10-2012-03.

REFERENCES

1. Atul Adya, Paramvir Bahl, Jitendra Padhye, Alec Wolman, and Lidong Zhou. 2004. A multi-radio unification protocol for IEEE 802.11 wireless networks. In Proceedings of the IEEE 1st International Conference on Broadnets Networks (BroadNets’04) . IEEE, Los Alamitos, CA, 210–217. https://doi.org/10.1109/BROADNETS.2004.8
2. Sam Anzaroot and Andrew McCallum. 2013. UMass Citation Field Extraction Dataset. Retrieved May 27, 2019 from <http://www.iesl.cs.umass.edu/data/data-umasscitationfield>
3. Martin A. Fischler and Robert C. Bolles. 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24, 6 (June 1981), 381–395. https://doi.org/10.1145/358669.358692
4. Chelsea Finn. 2018. Learning to Learn with Gradients. PhD Thesis, EECS Department, University of Berkeley.
5. Jon M. Kleinberg. 1999. Authoritative sources in a hyperlinked environment. J. ACM 46, 5 (September 1999), 604–632. https://doi.org/10.1145/324133.324140
6. Matthew Van Gundy, Davide Balzarotti, and Giovanni Vigna. 2007. Catch me, if you can: Evading network signatures with web-based polymorphic worms. In Proceedings of the first USENIX workshop on Offensive Technologies (WOOT ’07) . USENIX Association, Berkley, CA, Article 7, 9 pages.
7. James W. Demmel, Yozo Hida, William Kahan, Xiaoye S. Li, Soni Mukherjee, and Jason Riedy. 2005. Error Bounds from Extra Precise Iterative Refinement. Technical Report No. UCB/CSD-04-1344. University of California, Berkeley.
8. David Harel. 1979. First-Order Dynamic Logic. Lecture Notes in Computer Science, Vol. 68. Springer-Verlag, New York, NY. <https://doi.org/10.1007/3-540-09237-4>
9. Jason Jerald. 2015. The VR Book: Human-Centered Design for Virtual Reality. Association for Computing Machinery and Morgan & Claypool.
10. Prokop, Emily. 2018. The Story Behind. Mango Publishing Group. Florida, USA.
11. R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
12. Brian K. Reid. 1980. A high-level approach to computer document formatting. In Proceedings of the 7th Annual Symposium on Principles of Programming Languages. ACM, New York, 24–31. <https://doi.org/10.1145/567446.567449>
13. John R. Smith and Shih-Fu Chang. 1997. VisualSEEk: a fully automated content-based image query system. In Proceedings of the fourth ACM international conference on Multimedia (MULTIMEDIA ’96). Association for Computing Machinery, New York, NY, USA, 87–98. https://doi.org/10.1145/244130.244151
14. TUG 2017. Institutional members of the LaTeX Users Group. Retrieved May 27, 2017 from <http://wwtug.org/instmem.html>
15. Alper Yilmaz, Omar Javed, and Mubarak Shah. 2006. Object tracking: A survey. ACM Comput. Surv. 38, 4 (December 2006), 13–es. https://doi.org/10.1145/1177352.1177355

Abdelghaffar, H., Kamel, S., and Duquenoy, P. (2010). “Studying E-Government Trust in Developing Nations: Case of University and Colleges Admissions and Services in Egypt”. Proceedings of the International Information Management Association Conference, Utrecht, The Netherlands.

Abdelsalam, H., Reddick, C., Gamal, S., and Al-shaar, A. (2013). “Social Media in Egyptian Government Websites: Presence, Usage, and Effectiveness”. Government Information Quarterly, 30(4): 406-416.

Baldassare, M. (2000). “California in the New Millennium: The Changing Social and Political Landscape”. Berkeley: University of California Press.

Bertot, J. and Grimes, J. (2012). “Promoting Transparency and Accountability through ICTs, Social Media, and Collaborative E-Government”. Transforming Government: People, Process and

A  APPENDICES

In the appendix section, three levels of Appendix headings are available.

A.1 General Guidelines (AppendixH2)

1. Save as you go and backup your file regularly.
2. Do not work on files that are saved in a cloud directory. To avoid problems such as MS Word crashing, please only work on files that are saved locally on your machine.
3. Equations should be created with the built-in Microsoft® Equation Editor included with your version of Word. (Please check the compatibility at <http://tinyurl.com/lzny753> for using MathType.)
4. Please save all files in DOCX format, as the DOC format is only supported for the Mac 2011 version.
5. Tables should be created with Word’s “Insert Table” tool and placed within your document. (Tables created with spaces or tabs will have problems being properly typeset. To ensure your table is published correctly, Word’s table tool must be used.)
6. Do not copy-and-paste elements into the submission document from Excel such as charts and tables.
7. Footnotes should be inserted using Word’s “Insert Footnote” feature.
8. Do not use Word’s “Insert Shape” function to create diagrams, etc.
9. Do not have references appear in a table/cells format as it will produce an error during the layout generation process.
10. MS Word does not consistently allow the original formatting to be modified in the text. In these cases, it is best to copy all the document’s text from the specific file and paste into a new MS Word document and then save it.
11. At times there are font problems such as “odd” stuff/junk characters that appear in the text, usually in the references. This can be caused by a variety of reasons such as copying-and-pasting from another file, file transfers, etc. Please review your text prior to submission to make sure it reads correctly.

A.1.1 Preparing Graphics (AppendixH3)

1. Accepted image file formats: TIFF (.tif), JPEG (.jpg).
2. Scalable vector formats (i.e., SVG, EPS and PS) are greatly preferred.
3. Application files (e.g., Corel Draw, MS Word, MS Excel, PPT, etc.) are NOT recommended.
4. Images created in Microsoft Word using text-box, shapes, clip-art are NOT recommended.
5. IMPORTANT: All fonts must be embedded in your figure files.
6. Set the correct orientation for each graphics file.

A.2 Placeholder Text

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Vulputate sapien nec sagittis aliquam. Malesuada fames ac turpis egestas sed tempus urna. Posuere sollicitudin aliquam ultrices sagittis orci. Consequat id porta nibh venenatis cras sed felis eget. Pellentesque eu tincidunt tortor aliquam nulla facilisi cras fermentum odio. Tincidunt nunc pulvinar sapien et ligula ullamcorper malesuada proin. Tincidunt lobortis feugiat vivamus at augue. Eget nunc lobortis mattis aliquam faucibus. Egestas diam in arcu cursus euismod quis.

Erat pellentesque adipiscing commodo elit at imperdiet. In hac habitasse platea dictumst quisque sagittis purus. At lectus urna duis convallis. Eu mi bibendum neque egestas congue. Est ullamcorper eget nulla facilisi etiam dignissim diam. Sed ullamcorper morbi tincidunt ornare massa eget. Aenean vel elit scelerisque mauris pellentesque. Ullamcorper dignissim cras tincidunt lobortis feugiat vivamus. Cras fermentum odio eu feugiat pretium nibh. Congue eu consequat ac felis donec et odio pellentesque diam. Velit sed ullamcorper morbi tincidunt ornare massa eget egestas. In metus vulputate eu scelerisque felis imperdiet proin fermentum leo. Nulla malesuada pellentesque elit eget gravida cum.

Nullam ac tortor vitae purus faucibus ornare suspendisse. Libero enim sed faucibus turpis in eu mi bibendum neque. Sodales ut etiam sit amet nisl purus. Egestas diam in arcu cursus. Aliquet porttitor lacus luctus accumsan tortor. Pharetra magna ac placerat vestibulum lectus. Sit amet mauris commodo quis imperdiet massa tincidunt. In nisl nisi scelerisque eu ultrices vitae auctor. Nisi vitae suscipit tellus mauris a diam. Dui vivamus arcu felis bibendum ut tristique. Laoreet suspendisse interdum consectetur libero id.

Enim eu turpis egestas pretium. Nulla aliquet enim tortor at auctor urna. Id aliquet risus feugiat in. Non enim praesent elementum facilisis leo. Integer feugiat scelerisque varius morbi enim nunc faucibus. Egestas dui id ornare arcu odio ut sem nulla pharetra. Montes nascetur ridiculus mus mauris. Orci dapibus ultrices in iaculis. Enim sed faucibus turpis in eu mi bibendum neque. Faucibus pulvinar elementum integer enim neque volutpat ac tincidunt vitae. Et ultrices neque ornare aenean euismod elementum. Et pharetra pharetra massa massa ultricies mi quis hendrerit dolor. Tempus iaculis urna id volutpat lacus laoreet non curabitur gravida. Est velit egestas dui id ornare arcu odio. Eu facilisis sed odio morbi quis commodo odio. Lectus vestibulum mattis ullamcorper velit sed ullamcorper morbi tincidunt.

Eu non diam phasellus vestibulum lorem sed risus ultricies. Convallis aenean et tortor at risus viverra adipiscing at. Mauris pellentesque pulvinar pellentesque habitant morbi. Elementum sagittis vitae et leo duis. Massa enim nec dui nunc. Nisl tincidunt eget nullam non nisi est sit amet. Amet nisl purus in mollis nunc sed id semper. Fermentum leo vel orci porta non pulvinar neque laoreet suspendisse. Diam vel quam elementum pulvinar etiam non quam. Sagittis orci a scelerisque purus semper eget. Aliquet porttitor lacus luctus accumsan tortor. Integer vitae justo eget magna fermentum iaculis eu non diam. Egestas pretium aenean pharetra magna ac. Cursus metus aliquam eleifend mi in nulla. Cursus mattis molestie a iaculis at erat pellentesque adipiscing. Pulvinar pellentesque habitant morbi tristique senectus. Gravida cum sociis natoque penatibus et magnis dis parturient montes. In aliquam sem fringilla ut. Ut consequat semper viverra nam libero justo laoreet. Pellentesque diam volutpat commodo sed egestas.

Ornare arcu odio ut sem nulla pharetra diam. Ut enim blandit volutpat maecenas volutpat blandit aliquam. Tempus iaculis urna id volutpat lacus. Nascetur ridiculus mus mauris vitae. Venenatis cras sed felis eget velit aliquet sagittis id. Laoreet non curabitur gravida arcu ac tortor dignissim convallis aenean. Maecenas ultricies mi eget mauris pharetra et ultrices neque ornare. Egestas purus viverra accumsan in nisl nisi scelerisque eu ultrices. Tempus urna et pharetra pharetra massa massa. Pulvinar neque laoreet suspendisse interdum consectetur libero id.

Nisl rhoncus mattis rhoncus urna neque viverra justo nec ultrices. Morbi quis commodo odio aenean sed adipiscing diam donec. Neque gravida in fermentum et. Scelerisque purus semper eget duis at tellus. Volutpat blandit aliquam etiam erat velit scelerisque in dictum non. Odio ut sem nulla pharetra diam sit. Sed pulvinar proin gravida hendrerit lectus a. Diam ut venenatis tellus in metus vulputate eu scelerisque. Id semper risus in hendrerit. Vel quam elementum pulvinar etiam. Amet aliquam id diam maecenas ultricies mi. Auctor elit sed vulputate mi sit amet. Orci dapibus ultrices in iaculis nunc. Sed vulputate odio ut enim blandit volutpat maecenas volutpat. Auctor urna nunc id cursus metus. Integer enim neque volutpat ac tincidunt vitae.

Scelerisque in dictum non consectetur a erat. Vel risus commodo viverra maecenas accumsan lacus vel facilisis volutpat. Dignissim sodales ut eu sem integer vitae justo eget magna. Nunc non blandit massa enim nec dui nunc mattis enim. Sed vulputate odio ut enim blandit volutpat maecenas. Ante in nibh mauris cursus. Donec pretium vulputate sapien nec sagittis aliquam malesuada. Eu volutpat odio facilisis mauris sit amet massa. Blandit turpis cursus in hac habitasse platea dictumst quisque. Donec enim diam vulputate ut pharetra sit.

Magna fringilla urna porttitor rhoncus dolor purus non. Fames ac turpis egestas integer eget. Mattis rhoncus urna neque viverra. Laoreet sit amet cursus sit amet dictum sit amet. Vel pretium lectus quam id leo in vitae turpis massa. Euismod lacinia at quis risus sed vulputate odio ut. Lorem dolor sed viverra ipsum. Viverra justo nec ultrices dui sapien. Aliquam nulla facilisi cras fermentum odio eu feugiat pretium. Adipiscing commodo elit at imperdiet dui accumsan sit amet nulla. Morbi leo urna molestie at elementum eu facilisis sed. Habitant morbi tristique senectus et netus et malesuada. Viverra ipsum nunc aliquet bibendum enim. Integer vitae justo eget magna fermentum. Tincidunt id aliquet risus feugiat. Mauris ultrices eros in cursus turpis. Amet venenatis urna cursus eget nunc. Nisl nisi scelerisque eu ultrices vitae.

Non pulvinar neque laoreet suspendisse interdum consectetur libero. Facilisis leo vel fringilla est ullamcorper eget nulla facilisi. Ipsum dolor sit amet consectetur adipiscing elit pellentesque. Risus quis varius quam quisque id. Bibendum arcu vitae elementum curabitur vitae. Vitae et leo duis ut diam quam nulla. Orci eu lobortis elementum nibh tellus molestie nunc non blandit. Arcu odio ut sem nulla pharetra diam sit amet. Quis vel eros donec ac odio. Est lorem ipsum dolor sit amet consectetur adipiscing