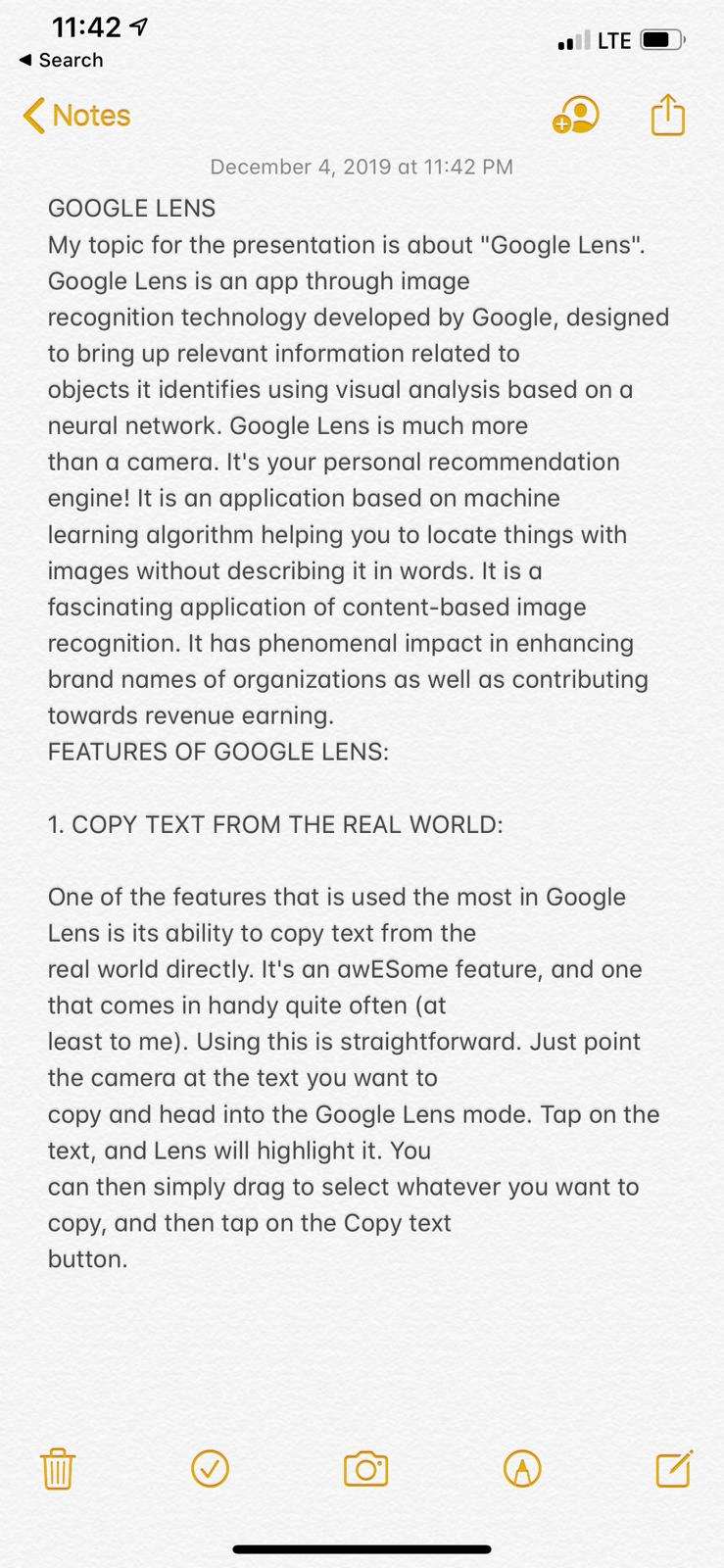
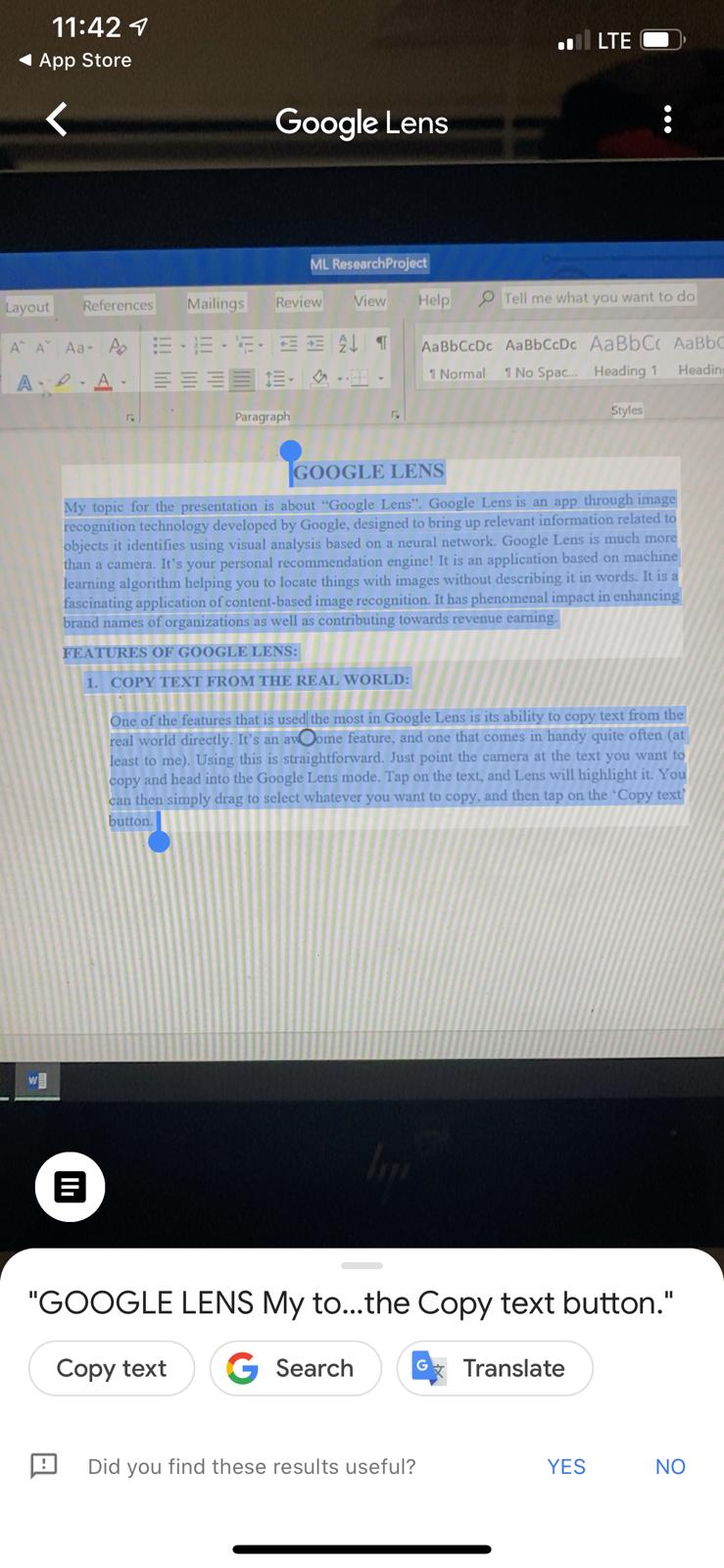
**GOOGLE LENS**

My topic for the presentation is about “Google Lens”. Google Lens is an app through [image recognition](https://en.wikipedia.org/wiki/Image_recognition) technology developed by [Google](https://en.wikipedia.org/wiki/Google), designed to bring up relevant information related to objects it identifies using visual analysis based on a neural network. Google Lens is much more than a camera. It’s your personal recommendation engine! It is an application based on machine learning algorithm helping you to locate things with images without describing it in words. It is a fascinating application of content-based image recognition. It has phenomenal impact in enhancing brand names of organizations as well as contributing towards revenue earning.

**FEATURES OF GOOGLE LENS:**

1. **COPY TEXT FROM THE REAL WORLD:**

One of the features that is used the most in Google Lens is its ability to copy text from the real world directly. It’s an awesome feature, and one that comes in handy quite often (at least to me). Using this is straightforward. Just point the camera at the text you want to copy and head into the Google Lens mode. Tap on the text, and Lens will highlight it. You can then simply drag to select whatever you want to copy, and then tap on the ‘Copy text’ button.



Here This is the picture of the MS Word that I was typing for the research project and I got all the data copied using the “Google lens” all I had to was hold my camera and press “copy text”.

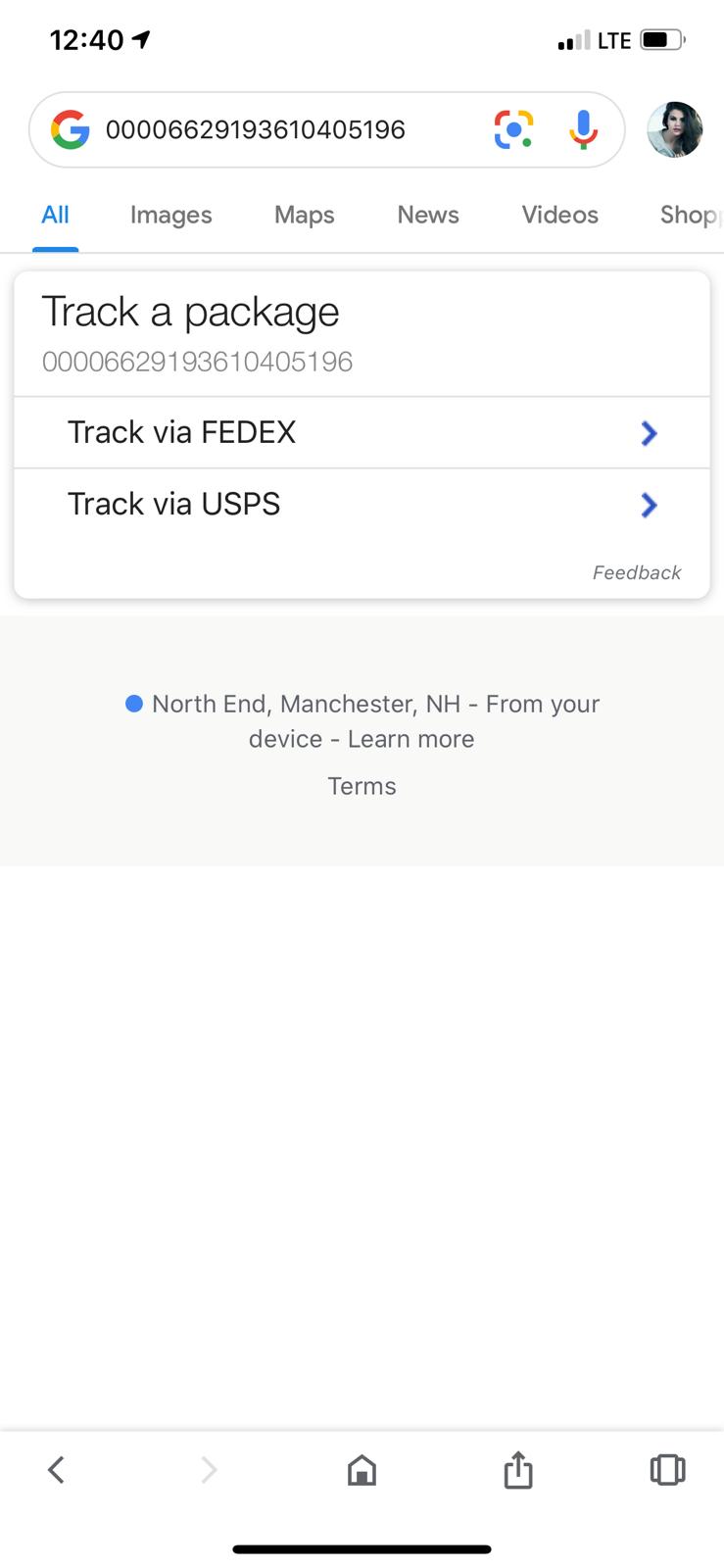
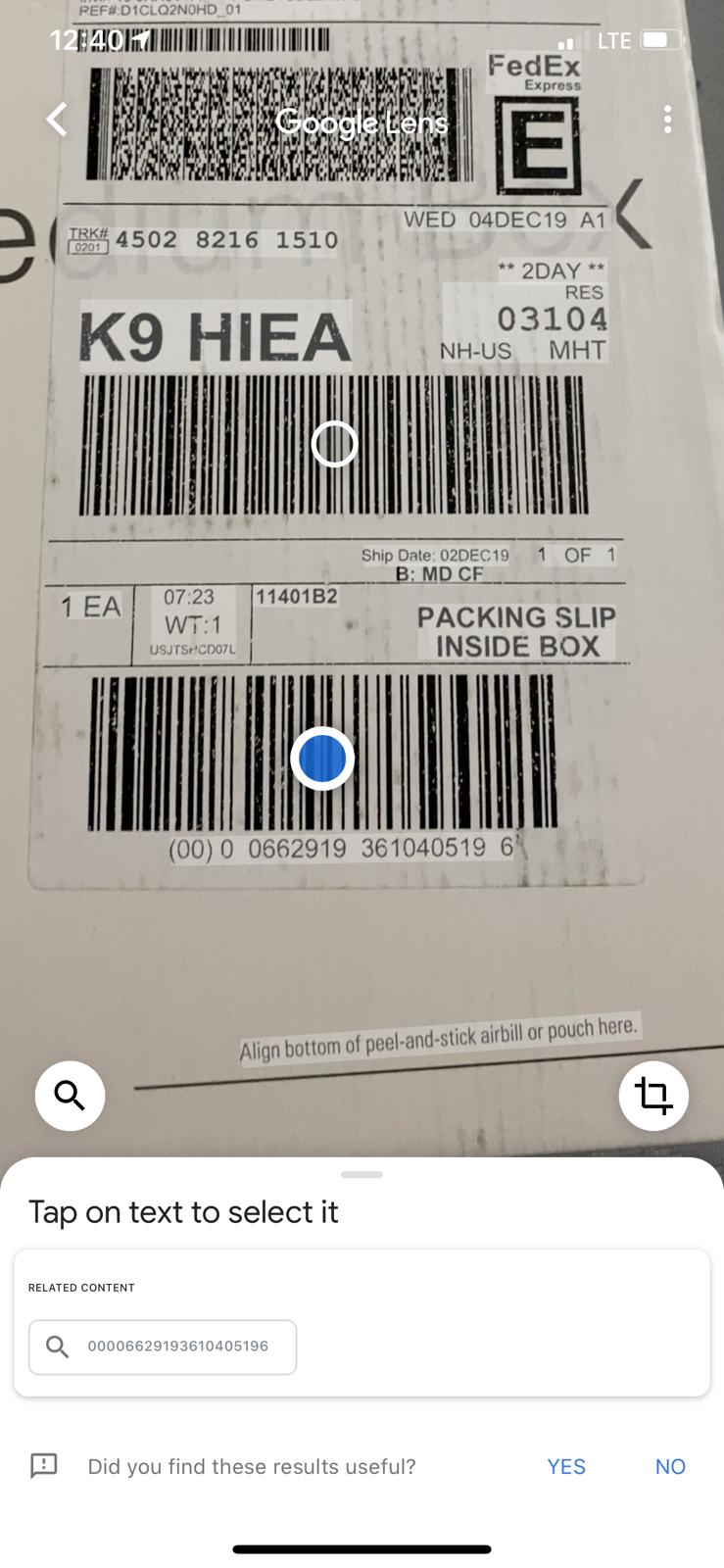
1. **WORD TRANSLATION:**

Position your camera in front of a book, newspaper or other surface that contains a word you wish to translate or find the meaning of. Make sure the camera is properly focused, so the word shows up clearly on the screen. Turn on Google Lens, and hold the camera steady for a few seconds until you see the word being highlighted in white on the camera screen. A dot will appear next to the word or text that is highlighted in white onscreen. Click on this dot to capture a picture of the highlighted text. You can now click on any word from the text to highlight it in blue, at which point you will see a list of options for interacting with the selected word.

****

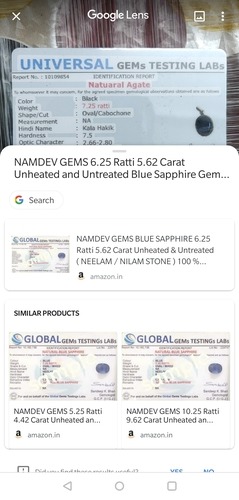
1. **TRACK PACKAGES BY SCANNING THE RECEIPTS:**

Position the camera lens in front of a package’s shipping details and turn on Google Lens. Wait until the details are highlighted, and then click on the dot that appears onscreen. You can simply scan a USPS, UPS or FEDEX receipt with Google lens, and it automatically shows the URL for the tracking page, so you can just tap on it and get tracking information about your package without having to go through the trouble of actually looking for the website, and then typing in your tracking number.



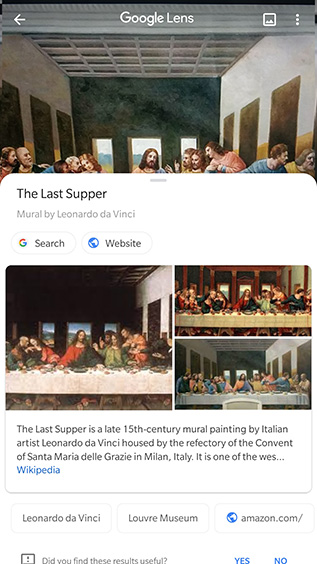
1. **ADD BUSINESS CONTACTS:**

Hold the business card in front of the camera lens and turn on Google Lens. Wait until the writing on the card is highlighted, and then click on the onscreen dot. The details of the card will appear at the bottom of the screen, with the option to go to the business’s website or save the details in your phone’s contact list. You can even use Google Lens to get directions to a business by scanning their business card itself. It only works for business that are listed on the Google Maps.



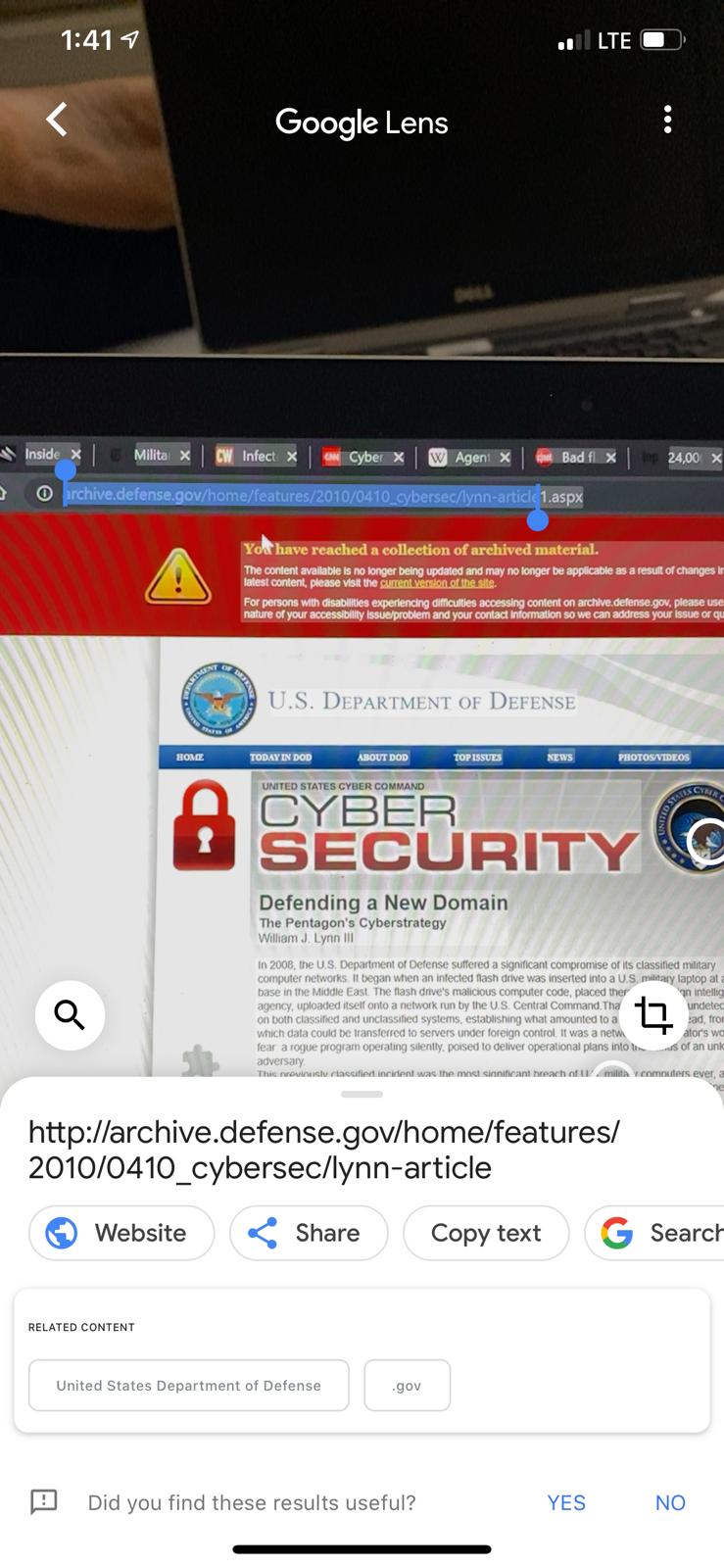
1. **DETECT FLOWERS, ANIMALS, PAINTINGS AND SHOPPING:**

Google Lens is also pretty good at detecting flowers, animals, famous paintings, and a lot of other stuff like landmarks. So, next time you’re out on a walk and you come across an unknown flower, try using Google Lens to figure out what it is. You can simply point your camera at the flower, **point your camera at the flower, animal (or object) to find what it is.** If that doesn’t work (sometimes it doesn’t), I find it easier to just**click a picture of it and use the Google Lens feature inside Google Photos.** Yeah, Lens is a little finicky right now, but we’ll give it some time to learn more about the world. That is, unless Google decides to send it the way of Hangouts, Google Plus.

****

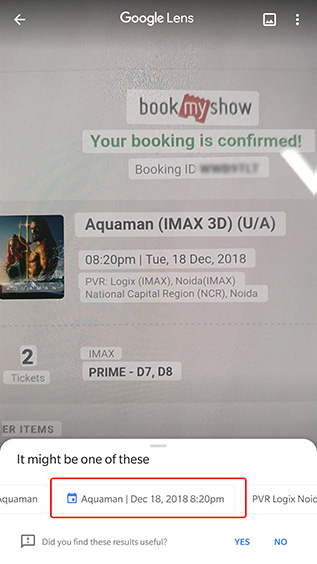
1. **SCAN URLs FROM PC TO OPEN IN MOBILE:**

Have you ever visited a website on your laptop and then wanted to visit the same website on your smartphone instead? With Google Lens, that’s easy. Simply launch Google Lens on your phone, point it at the URL bar on your browser, and tap on the URL. Google will **auto-detect the URL text and will give you the option to ‘go’ to that address in your smartphone.**

****

1. **ADDING EVENTS TO THE CALENDER:**

Google Lens can help you keep a track of events you need to attend, or that you bought tickets to. You don’t have to manually enter them into your calendar anymore. With Google Lens, you can **simply scan the ticket to a movie you booked tickets to, or an event you’re attending, and it’ll automatically suggest adding the event to your calendar.** It’s a handy feature that can save you the hassle of last-minute reminders for that date with your SO, or anything else that’s important to you. I’ve used it a bunch of times to scan movie tickets and it works well.

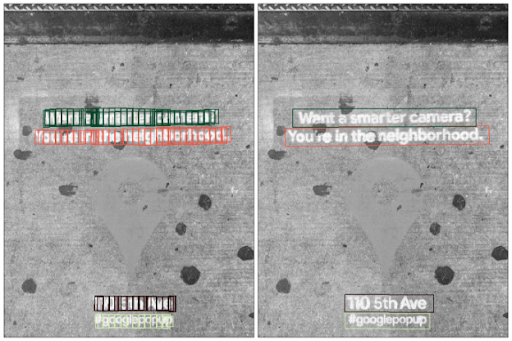
****

**ALGORITHMS BEHIND THE GOOGLE LENS:**

With the enhancement of machine learning techniques, especially in the domain of image processing and NLP, Google Lens has scaled to new heights. Here we look at a few algorithmic based solutions that power up Google Lens:

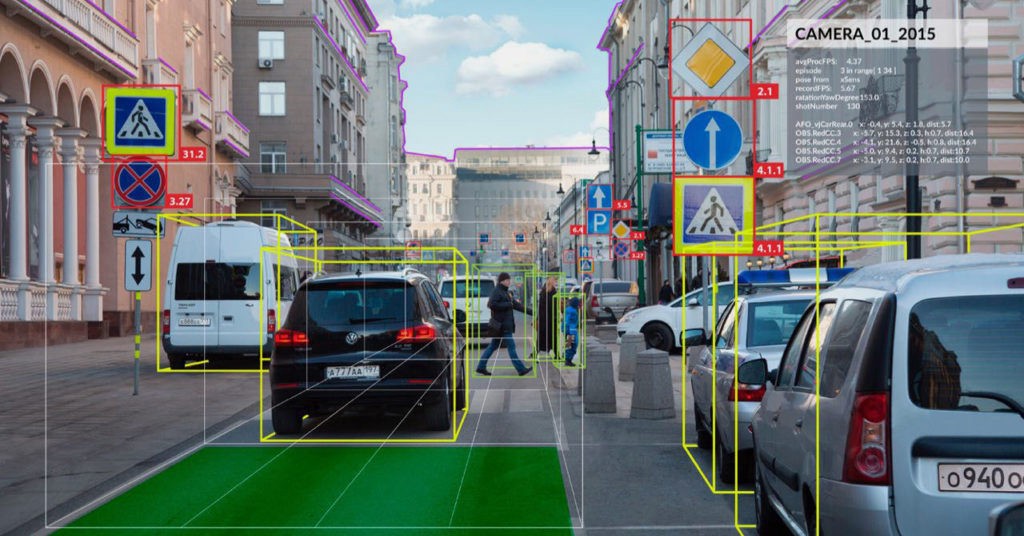
Lens uses computer vision, machine learning and Google’s Knowledge Graph to let people turn the things they see in the real world into a visual search box, enabling them to identify objects like plants and animals, or to copy and paste text from the real world into their phone.

### REGION PROPOSAL NETWORK:

[](https://www.analyticsindiamag.com/wp-content/uploads/2019/09/5-1.png)

After Google Lens captures an image, it needs to make sense of the shapes and letters. This is vital for text recognition tasks. So, Optical character recognition (OCR) utilizes a [region proposal network](https://arxiv.org/pdf/1506.01497.pdf) (RPN) to detect character level bounding boxes that can be merged into lines for text recognition.

RPN is a fully convolutional network that simultaneously predicts object bounds and objectness scores at each position. RPN is trained to generate high-quality region proposals, which are used by Fast R-CNN for detection. In short, its component tells the unified network where to look.



### KNOWLEDGE GRAPHS:

[](https://www.analyticsindiamag.com/wp-content/uploads/2019/09/3-1.png)

On the left there is an image with bounding box around recognized text. The raw optical character recognition (OCR) output from this image reads, “Cise is beauti640”.  By applying Knowledge Graph in addition to context from nearby words, Lens in Google Go recognizes the words, “life is beautiful”, as can be seen on the right.

The images captured by Google Lens may include sources such as signage, handwriting or documents, a slew of additional challenges can arise. The text can be obscured, stylized or blurry and it can cause the model to misunderstand words. To improve word accuracy, Lens utilizes the Knowledge Graph to provide contextual clues, such as whether a word is likely a proper noun and should not be spell-corrected and other such details.

### CONVOLUTIONAL NEURAL NETWORKS(CNNS):

The advent of large datasets and compute resources made convolution neural networks (CNNs) the backbone for many computer vision applications. The field of deep learning has in turn largely shifted toward the design of architectures of CNNs for improving the performance on image recognition.

Lens uses CNNs to detect coherent text blocks like columns, or text in a consistent style or color. And then, within each block, it uses signals like text-alignment, language, and the geometric relationship of the paragraphs to determine their final reading order.

All these steps, from script detection and direction identification to text recognition, are performed by separable convolutional neural networks (CNNs) with an additional quantized long short-term memory (LSTM) network. And the models are trained on data from a variety of sources, ranging from Captcha to scanned images from [Google Books](https://books.google.com/).

### NEURAL MACHINE TRANSLATION ALGORITHMS:

To provide users with the most helpful information, translations must be both accurate and contextual. Lens uses Google Translates [neural machine translation (NMT) algorithms](https://ai.googleblog.com/2016/09/a-neural-network-for-machine.html), to translate entire sentences at a time, rather than going word-by-word, in order to preserve proper grammar and diction.

For the translation to be most useful, it needs to be placed in the context of the original text.

For example, German sentences tend to be longer than English ones. To accomplish this seamless overlay, Lens redistributes the translation into lines of similar length, and chooses an appropriate font size to match. It also matches the color of the translation and its background with the original text using a heuristic that assumes the background and the text differ in luminosity, and that the background takes up most of the space.

**DEEPMIND WAVENET:**

[WaveNet](https://deepmind.com/blog/article/wavenet-generative-model-raw-audio) directly models the raw waveform of the audio signal, one sample at a time.  It is a fully convolutional neural network, where the convolutional layers have various dilation factors that allow its receptive field to grow exponentially with depth and cover thousands of timesteps.

The most helpful way with Lens in Google Go is reading the text aloud. For High-fidelity audio, google applies machine learning to disambiguate and detected entities such as dates, phone numbers and addresses, and uses that to generate realistic speech based on DeepMind’s WaveNet.