

Science AMA Series: I am Guangda Li, PhD in Media Computing and Co-Founder and CTO of ViSenze, a company developing visual search and image recognition through deep learning and computer vision.

GuangdaLi¹*andr/ScienceAMAs*¹

¹Affiliation not available

April 17, 2023

Abstract

Hi Reddit, The company originates from a spin-off from NExT, a leading research centre jointly established between National University of Singapore (ranked 22nd in the world) and Tsinghua University of China (ranked 47th in the world). The spin-off happened in 2012 and since then we have secured series A funding and well-known customers like Flipkart, Rakuten, Zalora from Rocket Internet and more. Deep learning is a very hot area at the moment. There were lots of developments in the past years that made it progress, but as of now the main evolution will come from the way it is implemented for specific applications. Visual technologies like visual search and image recognition are some of these specialisations that require not only a great use of deep learning and computer vision but also good industry knowledge for the verticals where it is applied. There is a huge talent crunch in this space and we need more and more engineers to consider a career in machine intelligence, deep learning and computer vision. I am here to answer questions regarding the real-world applications for deep learning and computer vision and what it takes to develop algorithms and infrastructure architectures from a research centre all the way to a company with established customers. I can reveal the industry potential and latest challenges, as well as why and how someone can develop a career in this space. AMA! Read more about my company: <https://visenze.com/> Explore live the product demo we have: <https://visenze.com/demo>

[REDDIT](#)

Science AMA Series: I am Guangda Li, PhD in Media Computing and Co-Founder and CTO of ViSense, a company developing visual search and image recognition through deep learning and computer vision.

GUANGDA_LI [R/SCIENCE](#)

ABSTRACT

Hi Reddit,

The company originates from a spin-off from NExT, a leading research centre jointly established between National University of Singapore (ranked 22nd in the world) and Tsinghua University of China (ranked 47th in the world). The spin-off happened in 2012 and since then we have secured series A funding and well-known customers like Flipkart, Rakuten, Zalora from Rocket Internet and more.

Deep learning is a very hot area at the moment. There were lots of developments in the past years that made it progress, but as of now the main evolution will come from the way it is implemented for specific applications. Visual technologies like visual search and image recognition are some of these specialisations that require not only a great use of deep learning and computer vision but also good industry knowledge for the verticals where it is applied.

There is a huge talent crunch in this space and we need more and more engineers to consider a career in machine intelligence, deep learning and computer vision.

I am here to answer questions regarding the real-world applications for deep learning and computer vision and what it takes to develop algorithms and infrastructure architectures from a research centre all the way to a company with established customers. I can reveal the industry potential and latest challenges, as well as why and how someone can develop a career in this space. AMA! Read more about my company: <https://visenze.com/> Explore live the product demo we have: <https://visenze.com/demo>

[READ REVIEWS](#)

[WRITE A REVIEW](#)

CORRESPONDENCE:

DATE RECEIVED:
October 30, 2015

DOI:
10.15200/winn.144612.20412

ARCHIVED:
October 29, 2015

CITATION:
Guangda_Li , r/Science ,
Science AMA Series: I am
Guangda Li, PhD in Media
Computing and Co-Founder
and CTO of ViSense, a
company developing visual

I'm currently working on a project about products like Google Glass and privacy concerns. One of the big pushbacks to Glass (in addition to the dorky look) was the concerns from the public about being monitored. Hacks that added facial recognition software added to these concerns. No one wants to feel as though they are being constantly monitored and that their faces are being scanned, identified, and recorded. Even though much of our everyday incorporates aspects of this from the facial recognition software employed at casinos to the programs platforms like Facebook utilize, Glass simply made it much more in your face and obvious. As people become more aware of facial recognition software there have been a variety of techniques promoted to fool it from how to do your makeup and hair to clothing you can wear. So I have two questions:

1. How is your industry dealing with public that purposefully try to fool image recognition software?
2. How do you manage the ethical concerns about privacy?

[firedrops](#)

search and image recognition through deep learning and computer vision., *The Winnower* 2:e144612.20412, 2015, DOI: [10.15200/winn.144612.20412](https://doi.org/10.15200/winn.144612.20412)

© et al. This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



At ViSenze we don't do 'face recognition'. Our majority of images are E-Commerce products and they are publicly available (even when they are provided by our API customers). We haven't encountered any privacy issue so far.

Dr. Li,

Can you elaborate on the how someone can most effectively get started in this field?

Thank you.

[TheEarthbound](#)

I assume the question is about the technology itself. I think there are plenty of online tutorial like CS231n: Convolutional Neural Networks for Visual Recognition from Stanford, and also open-source packages like Caffe and Torch which you can play with.

Do you think Computer Vision is based on sound assumptions? There's evidence that biological eyes behave more like change detectors, rather than something that provides a series of snapshots like you get out of digital imaging devices -- whether that's a grid of RGB pixels, or a RGB-D cloud. I know we've come a long way with computer vision, and I think Deep Learning is a very well suited technique to apply to it in terms of learning hierarchies of features, but do you think it's possible to replicate the capabilities of natural vision in an artificial system (by solving the same problems), given the underlying representation of what is observed is so different given our current technology? You're always constrained by that, right? Garbage-in, garbage-out, or more generally your ML algorithm is always bottlenecked by how good your input data is. Do you think our current input forms are good enough?

[mindrelay](#)

I agree with you. The current approach of doing computer vision research may be still in 'shallow' level if we look back after 20 years. But I believe that the fundamental technology will evolve and probably a good artificial system can replicate the human vision capacity in another 5-10 years. When I started my PhD study 8 years ago, I had never heard about deep learning but now everybody feels that this approach will probably achieve the ultimate AI goal.

But in my opinion, even though there are more and more research publications about unsupervised-based learning, the most efficient way in real industry application is still about getting high quality training data to improve the performance from the academic research.

Video sequence as input for training algorithm may be another angle to look at this issue as the information in video is so rich and the spatial-temporal information could be helpful in some way. But it is still too early to see some promising result.

If we talk about applied R&D like what ViSenze is doing now, our approach is that we we don't just focus on deep learning alone, but we keep focusing on our domain improvement: fundamental image processing and computer vision techniques for image and video. This builds tremendous performance improvement in our capability.

Thanks for the AMA! This might be a step outside of your normal work, but I am curious if you have any opinions on the application of deep learning to biological data, specifically to neuroimaging data. Full disclosure, I do not know too much about deep learning, so this might be naive.

Do you think deep learning can be more useful than supervised techniques when analyzing these types

of biological systems - which may be very noisy and of a structure that we do not fully understand?

[marsyred](#)

biological I think it is possible though I am not an expert in neuroimaging domain. But as far as I know, deep learning based method has won the top performance for most of typical computer vision tasks like recognition, segmentation, and tracking etc.

Recently, A company called Enlitic also uses deep learning to help radiologists to understand medical images better. I think it is also a sign that deep learning can be used for more domains.

A friend once told me that sentient machine intelligence isn't far off from where we stand and it's all about databasing and information stacking. How true is this information relating to machine intelligence that you work upon?

[freakshow9009](#)

In terms of information stacking, I think it is quite true. Basically, all type of data analytic and data modeling are about information stacking in some form.

How do you build a team at the moment given the particular skills needed for the engineers? If memory serves me correctly, most job openings that require knowledge of machine learning and, to a lesser extent, computer vision also require that the candidate have a master's or doctorate in computer science. But it takes time to develop a pool of candidates with that type and level of education. So how do you build a team in the meantime?

[midasgoldentouch](#)

It is quite true that most of job openings in this domain require knowledge and background both in machine learning and computer vision. And usually it requires thousands of hours of study and practice to be a 'master' in these fields. So that is the reason for campus hiring - we prefer students with advanced degrees at least.

For any high-tech company, building of the talents pipeline is very critical and it requires continuous effort. We try to find the good candidates even 1-2 years before they graduate. We engage with them and try to bring them into an internship program or collaborate with them.