Connecting relevant video content to audiences
Hello, We’re Vilynx

✓ We've developed **machine learning technology** to drive video discovery and engagement.

✓ **US company**, headquarters in Palo Alto, offices in New York and Barcelona (Spain).

✓ Top ML/DL team with **more than 200 patents** among everyone

✓ **Funding**: $2.6M seed Capital.
Millions of video are recorded every sec.

Millions of video are uploaded to the cloud every sec.

Video creation growing exponentially due to live feeds.
How any of us can navigate through all those hours of video?

- **Visually:** Static picture/thumbnail.
  - A static image cannot describe the most relevant moment of the video.

- **Search & Recommendation engine:** key words of the video/metadata & audience preference
  - No metadata from trending conversations in social media/web.
  - No metadata associated to relevant parts inside the video.
  - Most metadata currently added manually – doesn’t scale
How do we improve video discoverability?

- **Previews** - Real time selection of the best moments of a video.

- **Metadata** - Real time auto tagging those moments with social media/web key words.

- **Personalization** - Real time selection based on each user preferences.
Deep Networks today: match images to tags/key words

Images are not “moments/actions”.

What does Machine Learning & Deep Learning do today?
How does ML/DL auto-tagging work today?

Current training sets are based on public databases of images and the keyword/ tags are manually selected.

Those key words don’t describe what happens in real time. Live sporting events, terrorist attacks, concerts…..

People don’t search for ‘tennis court’ or ‘tennis player’, people look for ‘Wimbledon 2016’ and ‘Andy Murray’.

Auto-tagging video requires training datasets based on key words people use to describe those videos, moments, actions.
Do current solutions learn from videos and social media in real time today?

No they don’t.

How can we train hours of video in real time? Too much GPU/CPU processing

*Real time content requires a new architecture for ML/DL.*
How much info is relevant within a video for a Deep Network?

The inverse Autocorrelation function show that most of time is zero, no new info but ..... There are the key moments
How to move our Deep Learning Networks from images to “video moments”?

A data training set of video moments that will include:
- Set of frames that describes it: 5 sec clips.
- Key words that describes that moment, including social media.

Update the training set in real time with new videos and new key words.

Train the network in real time to be able to auto-tag new videos in real time.
We are creating a data set for DL training with 5 to 7 sec clips of videos and key words from social media / web

Lionel Messi Amazing Second Goal
Barcelona vs Bayern Munich score
2 - 0  May 06 2015 Champions League

James Harden final basket Rockets versus New York Nets
Here’s how it works.

- We ingest customer videos and the contextual information around it.

- We then take cues from around the Web and social networks.

- This combined input is fed to the most advanced convolutional deep neural network in the industry.

- Output are video previews optimized to engage your audience and rich metadata that can further drive your video content.
DL Through a Distributed Multi Convolutional Neural Networks

- Distributed network architecture
- Multiple Networks rather than a single huge network
- Real Time Training
- 80% accuracy on selecting frames per video

Taking advantage of all current high-performance HW, servers/GPUs
Use Video Processing Algorithms to find the relevant parts of the videos, taking into account:

- Image recognition (objects, people, places)
- Movement, colors, diversity
- Audience preferences
- Trending topics

Algorithms trained using ML (audience behavior analytics, individual scores per groups of videos, etc.)

More than 400k videos processed
300M summaries served
1.2B audience data points.

98% accuracy to find the relevant parts of the video
CTR increase between 50% to 500% (customer validated)
DL algorithms to extract video keywords

Our training DB learns from any online video and ads this information to the keyword selection

Reaching to 90% top-1 accuracy on keyword prediction

- Title
- Meta-Keywords
- Social Trends
- Full text
- Video URL
Huge database of tagged videos

* 400k videos with
* 20k unique tags clustered into
* 100 different clusters.

This results in
50M tagged images,
2.5M of video moments/actions
And growing every day
Thank You

Juan Carlos Riveiro
Chief Executive Officer
jc@vilynx.com

Palo Alto • New York City • Barcelona