DEEP LEARNING FOR SLOW MOTION VIDEO

Thomas True, Senior Applied Engineer for Professional Video
Andrew Page, Director Advanced Technologies
Deep Learning
WHAT IS DEEP LEARNING?
The differences between AI, ML & DL

Artificial Intelligence (AI): general coverall for machines doing interesting things

Machine Learning (ML): computers complete tasks without explicit programming

Neural Networks (NN): one technique to achieve ML

Deep Learning (DL): adds “hidden layers” to Neural Networks to solve complex problems

Great explanation: https://goo.gl/hkayWG
DEEP LEARNING APPLICATION DEVELOPMENT

TRAINING
Learning a new capability from existing data

INFEERENCE
Applying this capability to new data

Untrained Neural Network Model

Deep Learning Framework

TRAINING DATASET

Trained Model
New Capability

App or Service Featuring Capability

NEW DATA

Trained Model Optimized for Performance
### WHAT PROBLEM ARE YOU SOLVING?
#### Defining the AL/DL Task

<table>
<thead>
<tr>
<th>INPUTS</th>
<th>QUESTION</th>
<th>AI/DL TASK</th>
<th>EXAMPLE OUTPUTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text Data</td>
<td>Is “it” present or not?</td>
<td>Detection</td>
<td>Object Detection</td>
</tr>
<tr>
<td>Images</td>
<td>What type of thing is “it”?</td>
<td>Classification</td>
<td>Object Identification (Labeling)</td>
</tr>
<tr>
<td>Video</td>
<td>To what extent is “it” present?</td>
<td>Segmentation</td>
<td>Feature Tracking</td>
</tr>
<tr>
<td>Audio</td>
<td>What is the likely outcome?</td>
<td>Prediction</td>
<td>Denoised Pixel Values</td>
</tr>
<tr>
<td></td>
<td>What will likely satisfy the objective?</td>
<td>Recommendation</td>
<td>Animation Pose Selection</td>
</tr>
<tr>
<td></td>
<td>What would be a new variant?</td>
<td>Generation</td>
<td>Texture Creation</td>
</tr>
</tbody>
</table>
WHAT IS “SUCCESSFUL” DEEP LEARNING?

Many things but not “all the things!”

What characteristics do successful deep learning applications share?

1. Define the problem; Make sure DL is appropriate
2. Collect a lot of the right data (labelled data is even better)
3. Ensure ROI; choose a large task
Slow Motion Video
Defining the Problem
MULTI-FRAME VARIABLE LENGTH INTERPOLATION

Interpolate an Arbitrary Number of Frames Between Two Input Frames

\[ T = 0 \]

\[ T = 1 \]

[Mahajan et al. SIGGRAPH’09]
SINGLE-FRAME INTERPOLATION
Interpolate Only the Single Frame Inbetween

$T = 0$  $T = 0.5$  $T = 1$

RECURSIVE SINGLE-FRAME INTERPOLATION

No Time Order for Real-Time Applications

\( T = 0 \)  \hspace{1cm} \text{Step 2}  \hspace{1cm} \text{Step 1}  \hspace{1cm} \text{Step 2}  \hspace{1cm} \( T = 1 \)
RECURSIVE SINGLE-FRAME INTERPOLATION

Inefficient for Arbitrary Number of Frames

\[ T = 0 \quad T = 1/3 \quad T = 2/3 \quad T = 1 \]
MULTI-FRAME VARIABLE LENGTH INTERPOLATION

Can Interpolate in Correct Time Order

\[ T = 0 \quad \text{and} \quad T = 1 \]
MULTI-FRAME VARIABLE LENGTH INTERPOLATION

Can Interpolate in Reverse Time Order

\[ T = 0 \]

\[ T = 1 \]
MULTI-FRAME VARIABLE LENGTH INTERPOLATION

Can Interpolate All at Once

\[ T = 0 \]

\[ T = 1 \]
MULTI-FRAME VARIABLE LENGTH INTERPOLATION
Can Interpolate Only Selected Frames

\[ T = 0 \quad T = 1/3 \quad T = 2/3 \quad T = 1 \]
Super SloMo
Super SloMo: High Quality Estimation of Multiple Intermediate Frames for Video Interpolation


SUPER SLOMO
Frame Synthesis Using Optical Flow

$T = 0$  \hspace{2cm} $T = t \in (0,1)$  \hspace{2cm} $T = 1$
SUPER SLOMO
Visibility Information

$T = 0$

$T = t$

$T = 1$

$\alpha = 1$

$\alpha = 0$
OPTICAL FLOW
Image Motion Correspondence

Input

Optical flow

Color key
Baker et al. IJCV’11

Data credit: Liu et al. CVPR’08
OPTICAL FLOW
CNNs Estimates Flow Between Input Frames
FlowNet, FlowNet2, SpyNet, PWC-Net...

input images → bi-direct flow

optical flow computation between inputs
SUPER SLOMO OPTICAL FLOW

Challenge: Flow from Unknown to Input

\[ T = 0 \quad \quad T = t \quad \quad T = 1 \]
SUPER SLOMO OPTICAL FLOW

Simplified 1D Setting

\[ T = 0 \quad \rightarrow \quad T = 1 \]

\[ F_{0 \rightarrow 1} \]

\[ F_{1 \rightarrow 0} \]
SUPER SLOMO OPTICAL FLOW

Goal: Compute Motion Vector that Combines Forward and Backward Motion

\[ F_{t \rightarrow 0} \]

\[ F_{t \rightarrow 1} \]
SUPER SLOMO OPTICAL FLOW

Solution: Borrow from Bi-Directional Flow

\[ T = 0 \quad T = t \quad T = 1 \]

\[ F_{t \rightarrow 0} \quad F_{1 \rightarrow 0} \quad tF_{1 \rightarrow 0} \]

Same spatial location
SUPER SLOWMO OPTICAL FLOW

Solution: Borrow from the Bi-Directional Optical Flow

\[ F_{0 \rightarrow 1} = (1 - t)F_{0 \rightarrow 1} \]
SUPER SLOWMO OPTICAL FLOW

Solution: Borrow from the Bi-Directional Optical Flow

\[ T = 0 \quad T = t \quad T = 1 \]

\[ F_{t \rightarrow 0} \quad F_{0 \rightarrow 1} \quad (t - 1)F_{0 \rightarrow 1} \]
SUPER SLOWMO OPTICAL FLOW

Bilinear Combination of Two Approximations

\[ T = 0 \quad T = t \quad T = 1 \]

\[ F_{t \rightarrow 0} \approx t(tF_{1 \rightarrow 0}) + (1 - t)(t - 1)F_{0 \rightarrow 1} \]
 SUPER SLOMO NETWORK ARCHITECTURE

Add Flow Approximation

\[ F_{t \rightarrow 0} \approx t(tF_{1 \rightarrow 0}) + (1 - t)(t - 1)F_{0 \rightarrow 1} \]

- Input images
- Bi-direct flow
- Flow approx

Deterministic, parameter-free

Optical flow computation between input
ASSUMPTION: SMOOTH MOTION FIELD

\[
T = 0 \quad T = t \quad T = 1
\]

\[
F_{0 \rightarrow 1}
\]

\[
F_{1 \rightarrow 0}
\]
OPTICAL FLOW
Motion Fields are Piecewise Smooth

Input

Ground truth optical flow

MPI-Sintel [Butler et al. ECCV’12]

HAMA [Liu et al. CVPR’08]

Middlebury [Baker et al. ICCV’07]
PROBLEM: MOTION BOUNDARIES

\[ T = 0 \quad T = t \quad T = 1 \]

Motion boundary

\[ F_{0 \rightarrow 1} \]
PROBLEM: MOTION BOUNDARIES

\[ T = 0 \quad T = t \quad T = 1 \]

Motion boundary

\[ F_{0 \rightarrow 1} \]

\[ F_{1 \rightarrow 0} \]
PROBLEM: MOTION BOUNDARIES

\[ T = 0 \quad T = t \quad T = 1 \]

\[ F_{t \rightarrow 0} \quad ? \]

\[ (t - 1)F_{0 \rightarrow 1} \]

\[ F_{1 \rightarrow 0} \]

\[ tF_{1 \rightarrow 0} \]
FLOW REFINEMENT

Flow refinement module

- input images
- flow approx
- warped images
- refined flow
SUPER SLOMO OPTICAL FLOW
Refinement Near Motion Boundaries

input

intermediate flow approximations

intermediate refined flow

difference
FLOW REFINEMENT AND VISIBILITY REASONING

Flow refinement module

- input images
- flow approx
- warped images
- refined flow
- visibility maps
- prediction
intermediate optical flow

warped input images

visibility maps

prediction $\hat{I}_t$

w/ visibility maps

$\hat{I}_t = \alpha g(I_0, F_{t\rightarrow 0}) + (1 - \alpha) g(I_1, F_{t\rightarrow 1})$

prediction w/o visibility maps
**PREDICTION USING FLOW AND VISIBILITY MAP**

\[
\hat{I}_t = \alpha g(I_0, F_{t\to 0}) + (1 - \alpha) g(I_1, F_{t\to 1})
\]

\(g\): warping function

**prediction w/ visibility maps**

**prediction w/o visibility maps**
OVERALL NETWORK ARCHITECTURE
Combination of Two CNNs

optical flow computation between input

arbitrary-time image synthesis

at each time step $t$
TRAINING DATASET

YouTube 240-fps

Adobe 240-fps*

*provided by Oliver Wang
GENERATIVE ADVERSARIAL NETWORK (GAN)

G tries to fool D

The Art Forger

Generator/Decoder Network: $f^\text{gen}_\phi(z_i)$

The Art Critic

Discriminator Network: $f^\text{dis}_\phi(x_i)$

True or Fake image
## PERFORMANCE

### Running Time (PyTorch)

<table>
<thead>
<tr>
<th>Training (6 V100)</th>
<th>Inference (1 V100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>~1000 videos, 0.3M frames: 10 days</td>
<td>7 1280*720 images: 0.36s</td>
</tr>
</tbody>
</table>
EVS
EVS AI ENGINE

Regular camera

Regular feed

AI Engine

Slow Motion feed

Artificial Intelligence hallucinated intermediate images
EVS AI ENGINE
Creating On-Demand Slow Motion

- LSM
- Commands
- Instructions are sent to the AI engine
  - Normal speed baseband video or file
- LTC, UMD, … information
- Preview
- AI ENGINE
  - IN = X VIDEO FRAMES
  - OUT = NX VIDEO FRAMES
- Slow motion video baseband feed or file
EVS AI ENGINE
Mimicking a Supermotion Camera

- **CAMERA**
  - Normal speed baseband video

- **AI ENGINE**
  - IN = X VIDEO FRAMES
  - OUT = NX VIDEO FRAMES

- **REPLAY SERVER**
  - Slow motion video in multiple phases

- **LSM**
  - Commands
NVIDIA NGX: DL FOR CREATIVE APPLICATIONS

Delivering Deep Learning Research Into Creative Applications

- The framework to make **NV DL research available in end user applications** through a common service interface

- DL Features: **SloMo, Up-Res, InFilling, DLAA** ...
  - Created, trained and updated by NVIDIA, will be available to any application running on an NGX GPU

- **NGX SDK** - ISVs can add these DL features to their application without going through the expense of developing and training [https://developer.nvidia.com/rtx/ngx](https://developer.nvidia.com/rtx/ngx)
Results
QUALITY ISSUES

Motion Blur

Objects move too fast

EXPOSURE TIME
SHARPENING TECHNIQUES
QUALITY ISSUES

Missing Information

SIGNIFICANTLY REDUCED BY DOMAIN SPECIFIC TRAINING

Objects move too fast

Missing information
QUALITY ISSUES

Missing Information
HOW DO WE IMPROVE THE QUALITY?

More Training

- More training data with variety of frame rates from 30 fps to 240 fps
- Application Specific Training (Sports, Drama, etc.)
- Post process filtering
Getting Started in DL
DEEP LEARNING CONTAINERS

Took learnings from scores of teams across NVIDIA and industry

Built containers to speed and ease deployment

Goals:

- Get engineers working on Deep Learning in minutes
- Ensure full GPU acceleration and latest optimizations on all frameworks
- Isolate frameworks (some don’t play nice)
- Share, collaborate, and test applications across different environments
VIRTUAL MACHINES VS. CONTAINERS

Motivations

Packaging mechanism for applications

► Consistent and reproducible deployment
► Lightweight with faster startup than VMs
► Logical isolation from other applications (at the OS level)

No first-class support for GPUs was available in runtimes
GPU SUPPORT IN DOCKER: 1.0

Open-source Project on GitHub since 2016
- >2MM downloads
- >7.5K stars
- >13 MM pulls of CUDA images from DockerHub

Enabled Various Use-cases
- NGC optimized containers from NVIDIA
- Adopted by major deep learning frameworks
NVIDIA GPU CLOUD
Deep Learning Across Platforms

Use across workstation, local servers and cloud

To learn more about NVIDIA GPU Cloud, visit: https://nvidia.com/cloud

To sign up, go to: https://nvidia.com/ngcsignup
NVIDIA GPU CLOUD
Downloadable Containers