MonoPerfCap:
Human Performance Capture from Monocular Video
(with voiceover)

Weipeng Xu¹ Avishek Chatterjee¹ Michael Zollhöfer¹ Helge Rhodin²
Dushyant Mehta¹ Hans-Peter Seidel¹ Christian Theobalt¹
¹Max Planck Institute for Informatics, Saarland Informatics Campus ²EPFL
What is Deep Learning?

Physics + Logic

Input ➔ Feature Extraction ➔ Decision Rules ➔ Output
1: Car
0: Not Car

Physics + Machine Learning

Input ➔ Feature Extraction ➔ Classifier ➔ Output
1: Car
0: Not Car

Deep Learning

Input ➔ Representation Learning ➔ Output
1: Car
0: Not Car
Machine Learning: Application independent

Deep Learning: Coupled to the application
Computer programming by Deep Learning

Instead of how to do a task:

```python
# Example code snippet

# Import necessary libraries
import numpy as np
import tensorflow as tf

# Define the model
input_layer = tf.keras.layers.Input(shape=(224, 224, 3))

# Convolutional layers
conv1 = tf.keras.layers.Conv2D(32, kernel_size=3, activation='relu')(input_layer)
conv2 = tf.keras.layers.Conv2D(64, kernel_size=3, activation='relu')(conv1)

# Pooling layer
pool = tf.keras.layers.MaxPooling2D(pool_size=(2, 2))(conv2)

# Flatten layer
flatten = tf.keras.layers.Flatten()(pool)

# Fully connected layer
output = tf.keras.layers.Dense(1000, activation='softmax')(flatten)

# Create the model
model = tf.keras.models.Model(inputs=input_layer, outputs=output)
```

Give examples of what to do:

- Person
- Sheep
- Dog
- Car
- Horse
Big Labels
Big Labels
Big Labels
Current research state-of-the-art

Reduce annotation effort:

Interpretability:

Efficiency:

Security against adversarial attacks:
Role of the university

DL revolution came from academia:

ImageNet Classification with Deep Convolutional Neural Networks

No products; generic solutions:

$$L = \mu \sum_{i=1}^{N} \|b_i - V y_i\|_2^2 + \nu \|V^T V - A\|_F^2 + \eta \|C - V\|_F^2$$

Knowledge powers innovation: