**Learning Visual Storylines with Skipping Recurrent Neural Networks**

Gunnar A. Sigurdsson, Xinlei Chen, Abhinav Gupta

- We learn the **visual and temporal** parts of concepts: Wedding, Christmas
- Learn in an **unsupervised** way with recurrent networks
- Our S-RNN learns the **common long-term latent stories**
- We train one model for each **concept** (e.g., Paris)

---

**Method**

- **Train RNN over all possible subsequences**
- Rewritten as sequential update equations for E and M steps
- Fits in the RNN pipeline: *Infer next variable, then backprop*
- Prediction is **softmax over future images**
  - Images are represented as their fc7 features

---

**Storylines**

- **Appearance**
- **Time**
- **Human Events**

**Summarization**

- **Find a short summary of a photo album**
- Inferring latent variables gives a summary
- Album can be summarized in various ways:

---

**Prediction**

- **Predicting what comes next**
- Using the model sequentially

---

**Graphs**

- LSTM
- S-RNN
- LSTMSub
- S-RNN
- K-Means
- Samples

---

**GitHub**

[github.com/gsig](https://github.com/gsig)