

**Dog running through grass**

Running

Through

Grass

Dog

Output Dense layer

(vocabulary size)

Dense Layer

Timesteps equal to maximum length set for caption

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

LSTM

Word Embeddings

Dense Layer

CNN (InceptionV3)

**Image**

**Captions**