# An attention based dual learning Approach for video Captioning

- existing video captioning Model.

    ① template based language Model. → (generated caption → robotic)

    ⑪ sequence learning based Model. → more flexible than template based
                                    syntatically.

- two stage

    ① encoder stage — CNN used (video to ex compact vector)

    ② decoder stage — RNN used.

- exposure bias may lead to error acumulation
    ↳ once a "bad" word is generated by model, this error will be propogated
      & accumulated as length of the sequence increases.

- two Model proposed.

    ① A caption generation Model.