Connectionist Temporal Classification For Text Recognition

Yan Xia

November 13, 2015

Text Recognition is Not New

Traditional Way

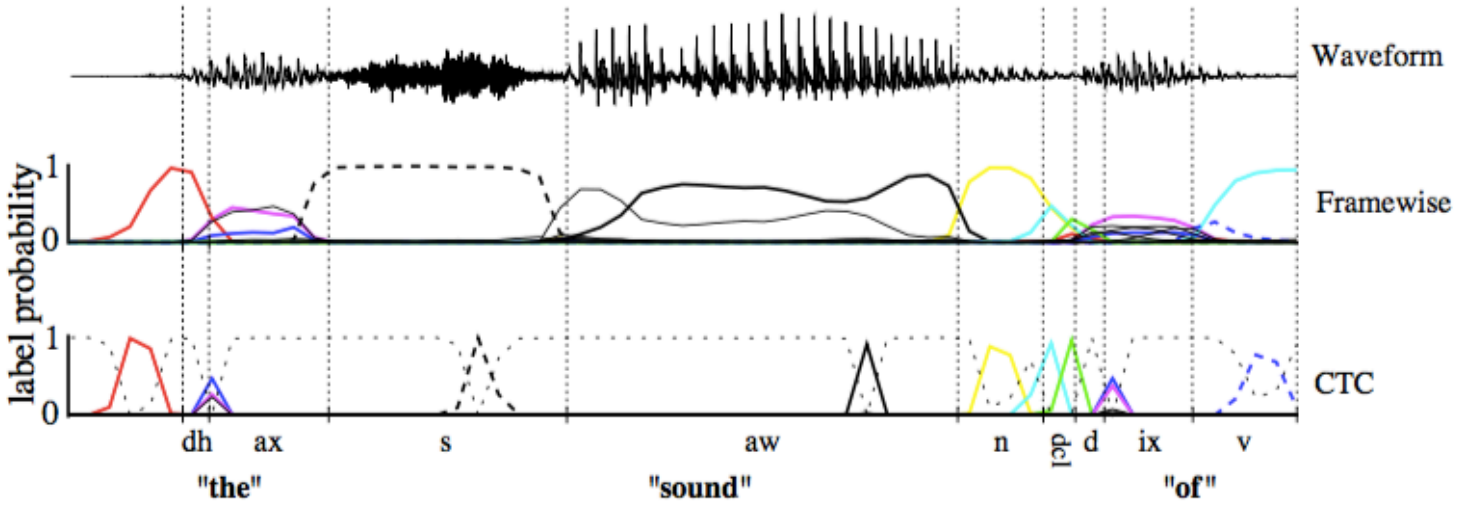
- char segmentation
  - binarization
  - connected region analysis
- char recognition
  - feature extraction
  - classification
- post-processing
  - text correction using language model
  - ...

Traditional Way

- Parameters in char segmentation changes a lot for different scenarios
- Require basically clean image background
- Need a lot of rules

Connectionist Temporal Classification (CTC)

- Alex Graves, Supervised Sequence Labelling with Recurrent Neural Networks, 2012



CTC

- Given a dict A and an input sequence x width length T, where x is consist by predictions y_k^t (k^{th} class at time t)
- Extend dict by $A' = A \cup \{blank\}$
- Conditional distribution over $\pi \in A'^T$

$$p(\pi|x) = \prod_{t=1}^{T} y_{\pi_t}^t$$

CTC

- Define a many-to-one function $\mathcal{F} : A'^T \mapsto A^{\leq T}$
- Eg: $\mathcal{F}(a - ab-) = \mathcal{F}(aaa - aabb-) = aab$
-

$$p(l|x) = \sum_{\pi \in \mathcal{F}^{-1}(l)} p(\pi|x)$$

CTC

- Now we have

$$p(\pi|x) = \prod_{t=1}^{T} y_{\pi_t}^t \tag{1}$$

$$p(l|x) = \sum_{\pi \in \mathcal{F}^{-1}(l)} p(\pi|x) \tag{2}$$

- Loss function (on training set S width label)

$$\mathcal{L}(S) = -\ln \prod p(l|x) = -\sum \ln p(l|x) \tag{3}$$

CTC

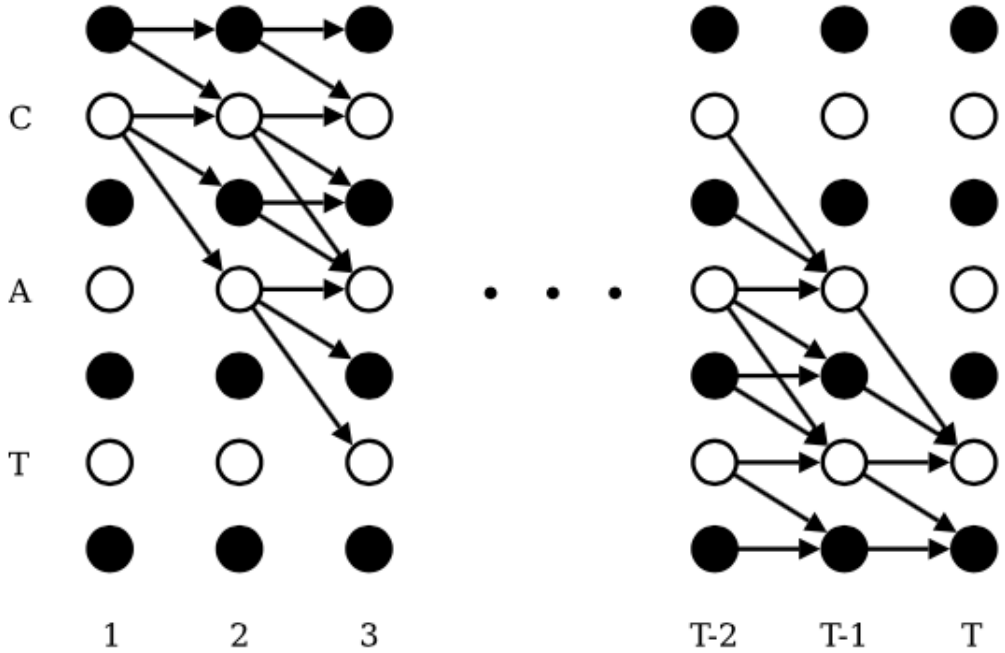
- DP. forward variable α:

$$\alpha(t,u) = \sum_{\pi \in V(t,u)} \prod_{i=1}^{t} y_{\pi_i}^{i} \tag{4}$$

$$p(l|x) = \alpha(T, U') + \alpha(T, U' - 1) \tag{5}$$

$$\alpha(t,u) = y_{l'_u}^{t} \sum_{i=f(u)}^{u} \alpha(t-1, i) \tag{6}$$

CTC

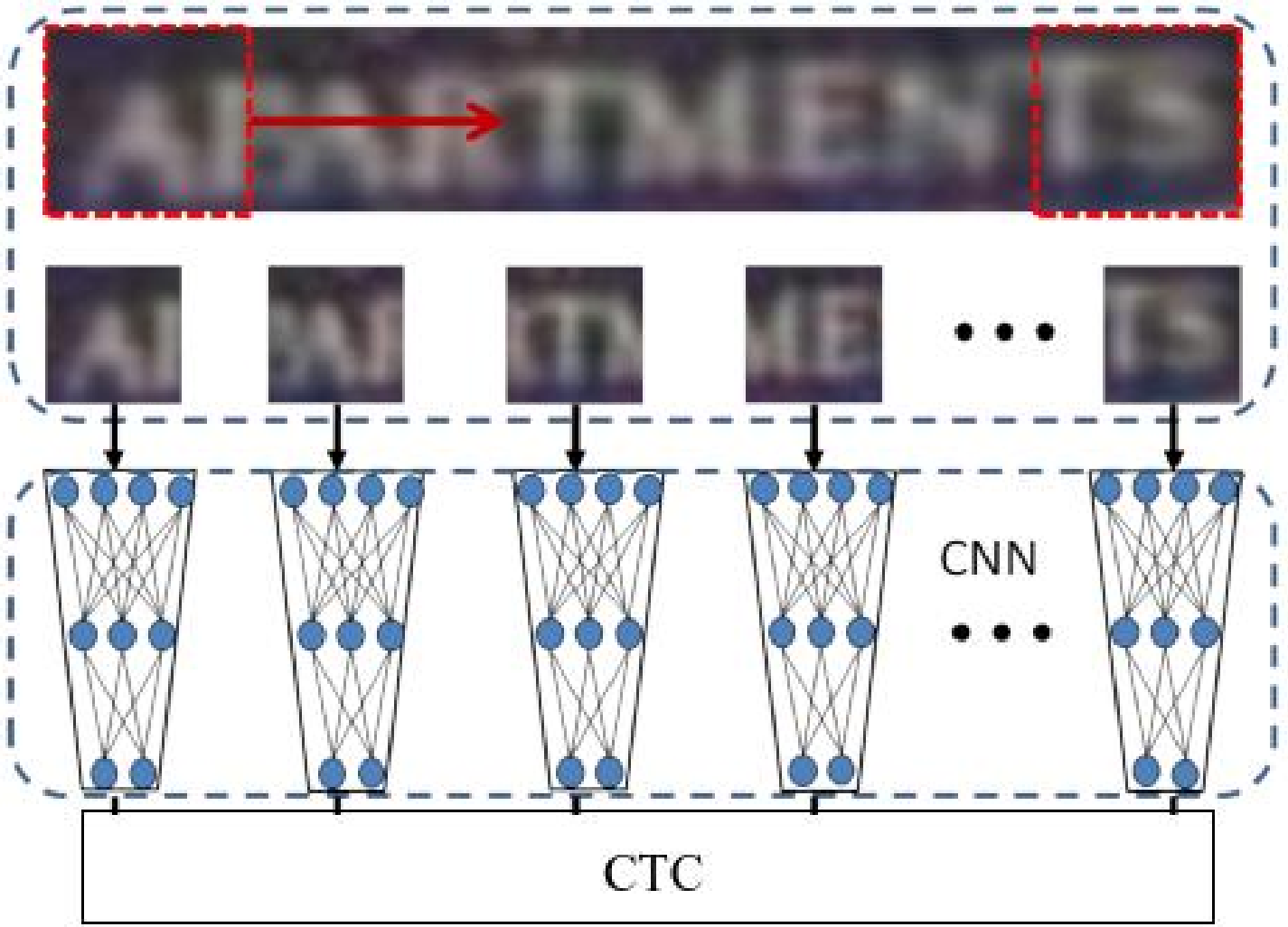


CTC

- The backward variable β is similar as α, and $\beta(t, u)$ defines the probability of all paths starting at $t+1$ that complete l when appended to $\alpha(t, u)$
- Finally, we have

$$p(l|x) = \sum_{u=1}^{|l'|} \alpha(t, u)\beta(t, u) \tag{7}$$

Text Recognition



Text Recognition


```
GENERAL INFO:
INFO 0
流水号: 338 (number)
INFO 1
上次累计积分:   466.10, (point)
INFO 2
本次积分:       1 (point)
INFO 3
time:   20h54m47s
INFO 4
date:   2015y07m23d
INFO 5
sum:    1.90
INFO 6
shop:   华润万家(crv
total info count:7
--------------------
GOODS INFO :
GOODS 0
name:   康师傅矿物质水1500ml
money:  1.90
count:  1
total goods count:1
```

Thanks!