

# 基于TCN的端对端车牌识别

## 既存方法

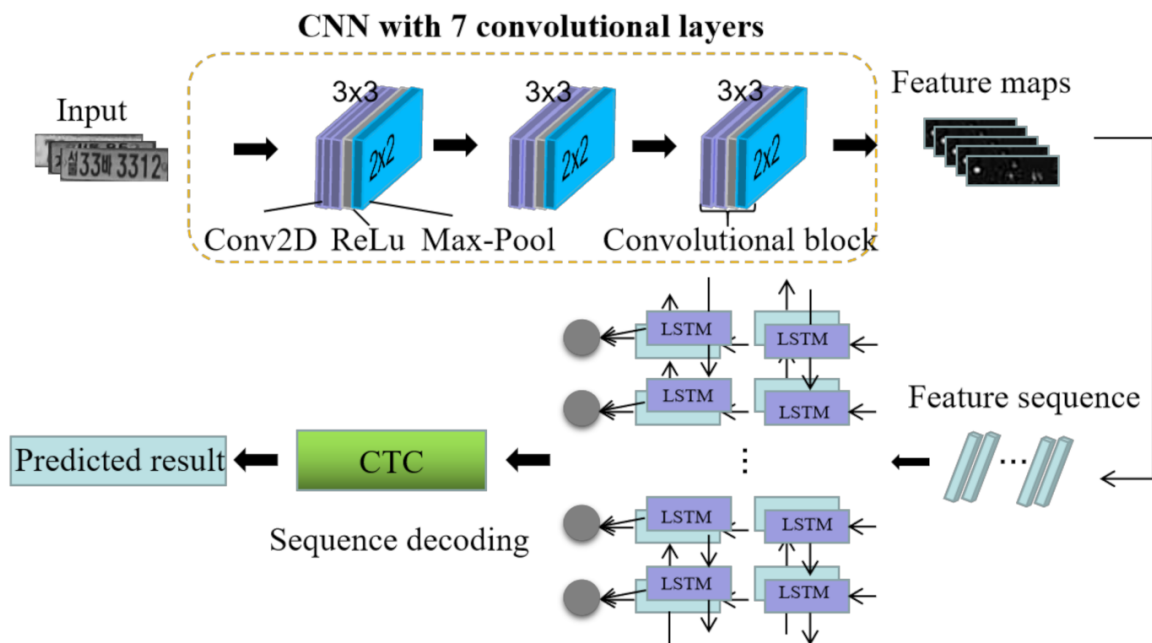
### 传统计算机视觉



### 存在的问题

过于依赖字符分割的准确性，字符分割的准确与否直接影响识别的结果。

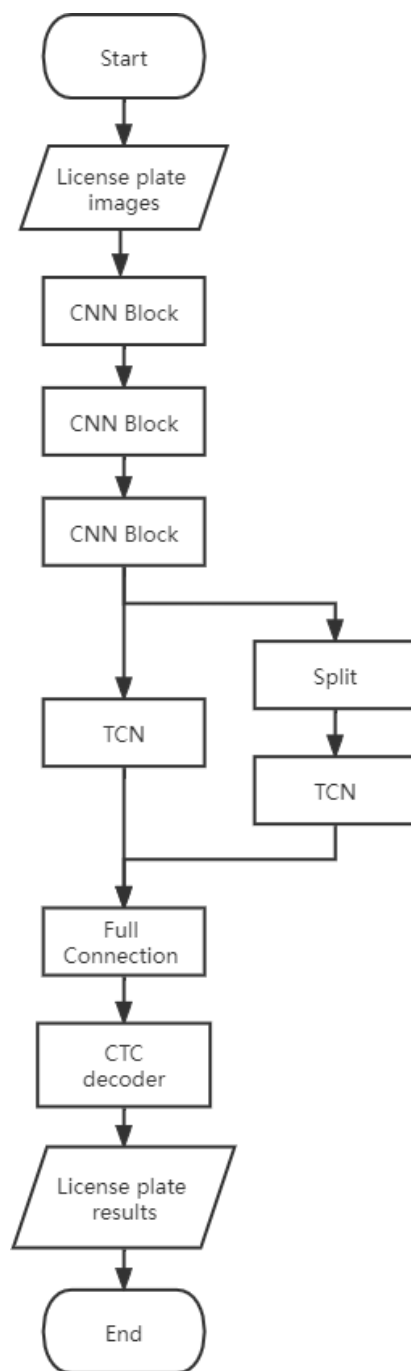
### CNN-LSTM 结构的基于深度学习的方法



## 存在的问题

LSTM结构较为复杂，在嵌入式设备部署中需要消耗大量的内存。

## 基于TCN 的车牌识别的方法



通过 **2D CNN** 提取特征，其中 **CNN Block** 主要结构如下

$$MaxPool(ReLU6(Conv(M))), ReLU6 = \min(6, \max(0, x))$$

对于提取到的特征，分割成两个部分，对上半部分进行TCN，

并且同样的将整个提取到的特征同样通过TCN，但这两个TCN的层数和每个filter的数量不同。

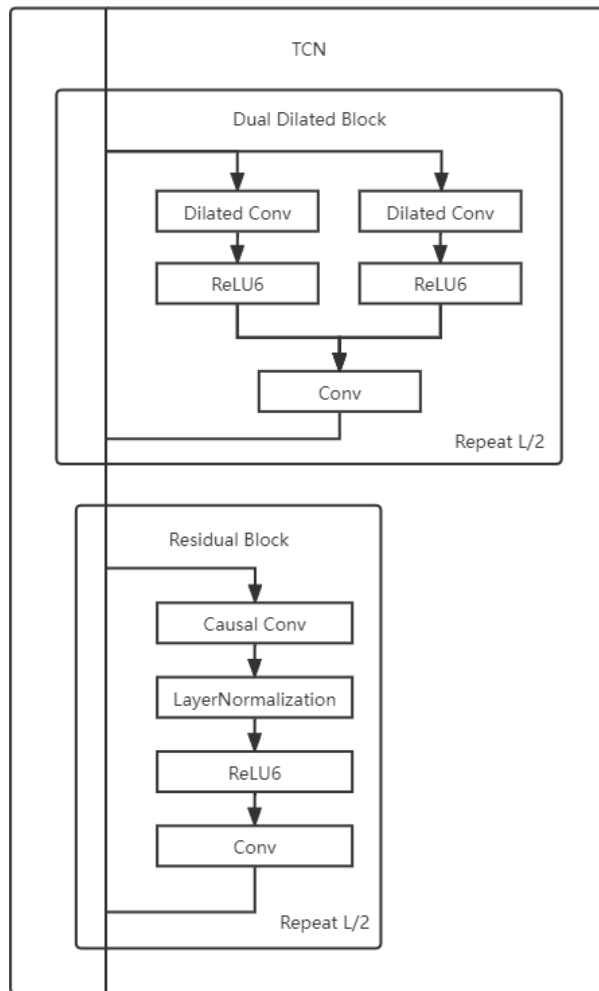
最后将这两个TCN提却的结果合并后并通过一个全连接层

实际运用的时候，对全连接层的输出利用 Greedy Search 进行解码

Greedy Search 公式如下

$$A^* = \arg \max_A \prod_{t=1}^T p_t(a_t|X)$$

## TCN结构



TCN 结构主要由 **Dual Dilated Block** 和 **Residual Block** 组成,

其中 **Dual Dilated Block** 结构如下

$$v_1 = \text{ReLU6}(\text{DilatedConv}(M, d = 2^{\frac{L}{2}-1-l}))$$

$$v_2 = \text{ReLU6}(\text{DilatedConv}(M, d = 2^l))$$

$$Y_1 = W[v_1, v_2] \in \mathbb{R}^n$$

$$Y_2 = W(M + Y_1)$$

其中 *DilatedConv* 公式如下

$$F(s) = (X *_d f)(s) = \sum_{i=0}^{k-1} f(i) \cdot \mathbf{x}_{s-d \cdot i}$$

**d** is dilation scale factor

其中 **Residual Block** 的结构如下

$$v = \text{ReLU6}(\text{LayerNormalization}(\text{CausalConv}(M, d = 2^l)))$$

$$Y = W(M + v)$$

// 可能不需要

最后添加CTC loss

$$p(Y|X) = \sum_{A \in A_{X,Y}} \prod_{t=1}^T p_t(a_t|X)$$