

TrackNet v1

NCTU

TrackNet: maximise $\partial L(x, y, t) / \partial t$

TSM: model $P(\text{Action} | f(t-k, t+k))$

channel 代表在不同時間下同一個 pixel 在某種的類別
 $L = \|\text{heatmap_pred} - \text{heatmap_gt}\|_1$

channel 代表不同證據 feature (抽像特徵, 較似 LHM 方式)
 $L = \text{CE}(\text{action_pred}, \text{action_gt})$

對比

TrackNet , Archana (BG sub)

資料

球和背景重疊 , 球被球員遮擋 , 球速太快有延長軌跡

軌跡模式

飛行 , 擊球 , 弹跳

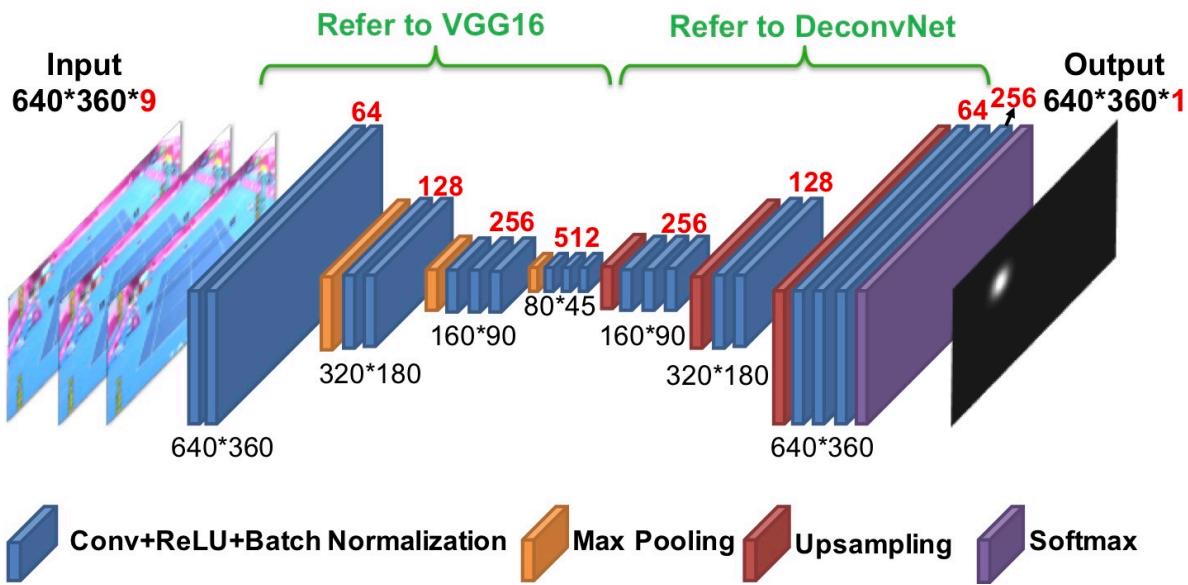
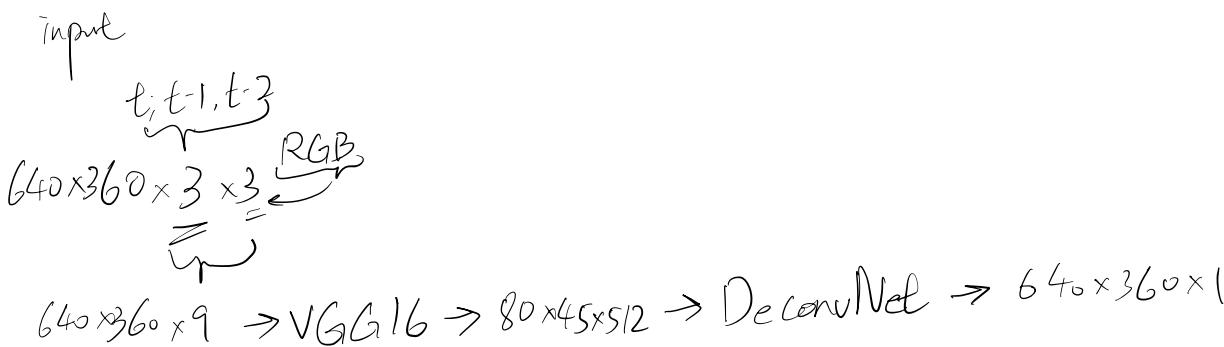
場地

草地 , 紅土 , 硬地

標記

FrameName , Visibility Class , X , Y

網絡架構



output heatmap 轉換做單向二元熱力圖，閾值為 128

$\geq 128 \rightarrow 255$

$< 128 \rightarrow 0$

\downarrow 二極化 heatmap

Circle hough \rightarrow 半徑法到點

\downarrow 有 附近

返回 圖片

TrackNet V2

Layer	Details	Output size
input	npy file	512 × 288 × 9
conv2d_1	$3 \times 3 \times 64$; relu; batch norm	512 × 288 × 64
conv2d_2	$3 \times 3 \times 64$; relu; batch norm	512 × 288 × 64
max_pooling_1	2×2 max pool; stride 2	256 × 144 × 64
conv2d_3	$3 \times 3 \times 128$; relu; batch norm	256 × 144 × 128
conv2d_4	$3 \times 3 \times 128$; relu; batch norm	256 × 144 × 128
max_pooling_2	2×2 max pool; stride 2	128 × 72 × 128
conv2d_5	$3 \times 3 \times 256$; relu; batch norm	128 × 72 × 256
conv2d_6	$3 \times 3 \times 256$; relu; batch norm	128 × 72 × 256
conv2d_7	$3 \times 3 \times 256$; relu; batch norm	128 × 72 × 256
max_pooling_3	2×2 max pool; stride 2	64 × 36 × 256
conv2d_8	$3 \times 3 \times 512$; relu; batch norm	64 × 36 × 512
conv2d_9	$3 \times 3 \times 512$; relu; batch norm	64 × 36 × 512
conv2d_10	$3 \times 3 \times 512$; relu; batch norm	64 × 36 × 512
up_sampling_1	2×2	128 × 72 × 512
concatenate_1	with conv2d_7; axis = 1	128 × 72 × 768
conv2d_11	$3 \times 3 \times 256$; relu; batch norm	128 × 72 × 256
conv2d_12	$3 \times 3 \times 256$; relu; batch norm	128 × 72 × 256
conv2d_13	$3 \times 3 \times 256$; relu; batch norm	128 × 72 × 256
up_sampling_2	2×2	256 × 144 × 256
concatenate_2	with conv2d_4; axis = 1	256 × 144 × 384
conv2d_14	$3 \times 3 \times 128$; relu; batch norm	256 × 144 × 128
conv2d_15	$3 \times 3 \times 128$; relu; batch norm	256 × 144 × 128
up_sampling_3	2×2	512 × 288 × 128
concatenate_3	with conv2d_2; axis = 1	512 × 288 × 192
conv2d_16	$3 \times 3 \times 64$; relu; batch norm	512 × 288 × 64
conv2d_17	$3 \times 3 \times 64$; relu; batch norm	512 × 288 × 64
conv2d_18	$1 \times 1 \times 3$; relu	512 × 288 × 3

TABLE I
TRACKNETV2 MODEL STRUCTURE.

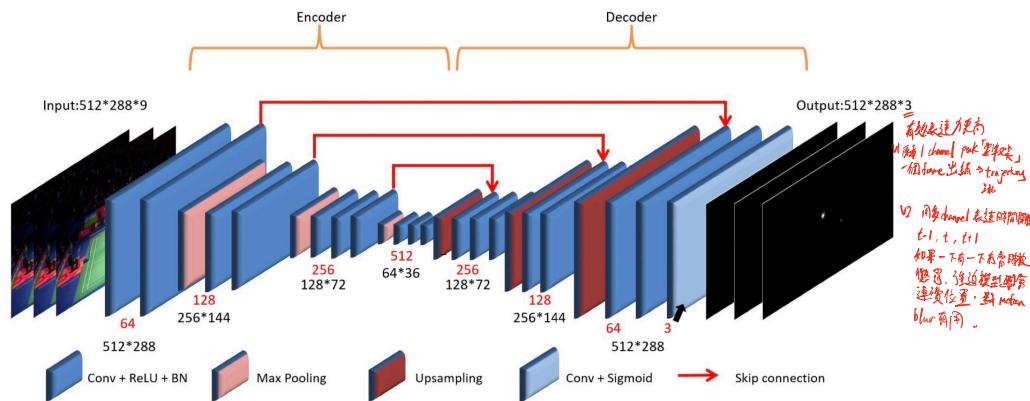


Fig. 2. The architecture of TrackNetV2. The encoder-decoder structure basically follows the design of TrackNet. Compared with the original TrackNet, the innovation includes the downsizing of input images, skip connections enlightened by U-net, concise heatmap representation, and MIMO network design.

V1 GT heatmaps 傷向 hard binary

Loss 對 notan blur frame 不穩定 (BCE)

decoder 上採樣太 aggressive

Output 只有 1 channel

BG 標本少，少 FP
data imbalance

V2 Gaussian heatmap 更平滑 (類似 soft label)

根據 size / resolution 調整

peak 當作一個 pixel

decoder 更 gradual

Output 3 channel

to BG frame, FP更敏感

class balance (2D vs 非2D)

warm-up

skip connection

FP 少，Trajectory jitter 少

Training 穩定，

TrackNet v3

input 中混入背景影像

Training 時使用 Mixup

羽球轨迹校正 (训练了羽球轨迹校正模块) 解决 物体遮蔽、视觉上遮挡、羽球飞出

