

Classificazione di Azioni Cestistiche mediante Tecniche di Deep Learning

Candidato: Francia Simone

Relatore: Prof. Simone Calderara

Correlatore: Dott. Fabio Lanzi

12 Aprile 2018, Modena

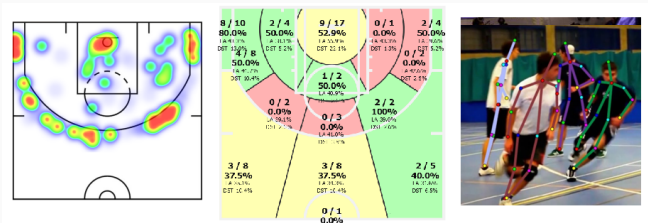
Università di Modena e Reggio Emilia

Corso di Laurea Magistrale in Ingegneria Informatica

Facoltà di Ingegneria "Enzo Ferrari"

Anno Accademico 2016-2017

- L'estrazione di informazioni e di indici prestazionali in ambito sportivo sta sempre più acquisendo importanza.
- Al crescere dei profitti ottenuti dallo sport, allo stesso modo le squadre sportive investono sempre maggiormente nella raccolta di statistiche.
- Tali statistiche possono essere utilizzate sia per il miglioramento di prestazioni della propria squadra, sia per raccogliere informazioni delle squadre avversarie.



Obiettivo: Classificare azioni personali in video di partite di basket

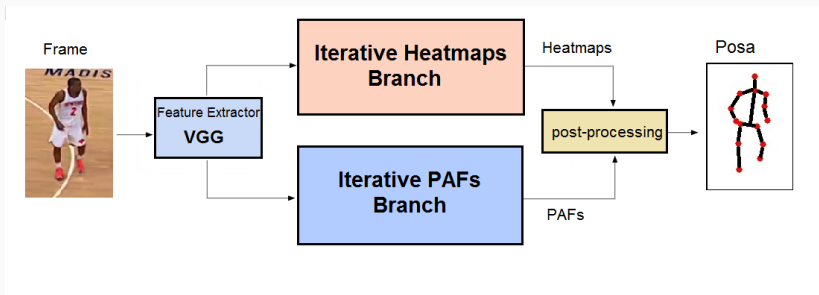
Fase 1: Creazione di Dataset di Azioni Cestistiche

- Dataset dei Giunti “jointsDataset” → parti del corpo
- Dataset delle Clip “clipsDataset” → finestra del giocatore

Fase 2: Applicazioni di Metodi di Deep Learning per la Classificazione ed Esperimenti

- Rete Completamente Connessa “jointFC” di tipo deep NN
- Rete “Convolutional 3D” o “C3D” di tipo deep NN
- Fusione jointFC e C3D nel modello “unionFC3D”

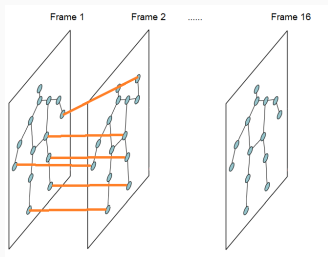
- Vari step di processing per creazione dei Dataset



- Convolutional Pose Machines utilizza metodi di Deep Learning
- Iterative Heatmaps Branch → Confidence Maps
- Iterative PAFs Branch → Campi Vettoriali
- Post-Processing(Heatmaps + PAFs) → Posa

2° step: Tracking di Pose tramite Nearest Neighbor

- Obiettivo: Fornire lo stesso identificatore di posa per la stessa persona tra frame contigui.

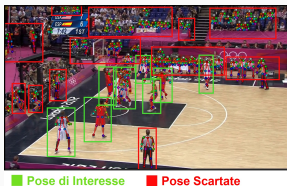


- Formula Nearest Neighbor Tracker applicato alle pose tra frame contigui:

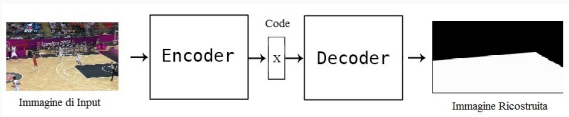
$$p_k(f+1) = \underset{p_j}{\operatorname{argmin}} \left\{ \sum_g^{\text{joints}} \sqrt{(g_{p_k}(f) - g_{p_j}(f+1))^2} \right\} \quad (1)$$

3° step: Segmentazione Binaria Campo Cestistico mediante Autoencoder e Filtraggio delle Pose

- Individuazione delle pose di interesse: Giocatori



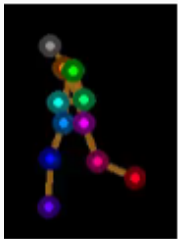
- Autoencoder per la Segmentazione Binaria del Campo



- Filtraggio delle pose tramite giunti “piedi”

4^o step: Creazione jointsDataset e clipsDataset

Dataset Giunti

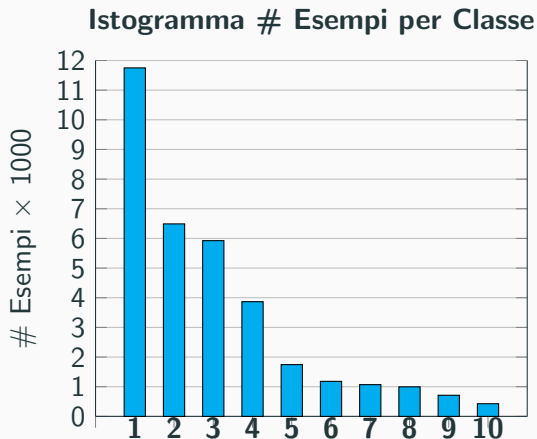


- 14 coordinate cartesiane (x,y) per 16 frame
- coordinate normalizzate in una finestra (176,128)

Dataset Clip



- clip di 16 frame in RGB
- bounded in (176,128)
- top-left,bottom-right mediante giunti



Classi di Azioni

1 → Camminata

2 → Nessuna Azione

3 → Corsa

4 → Difesa

5 → Palleggio

6 → Palla in Mano

7 → Stopzata

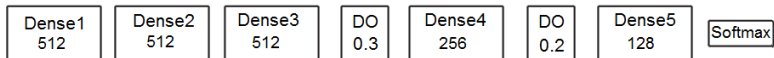
8 → Blocco

9 → Stopzata

10 → Tiro

- Video Partite in Full-HD (1920x1080 pixels)
- 15 partite dalla durata di 1.5 ore ciascuna
- Eliminazione di inquadrature e primi piani con Tecniche di VA

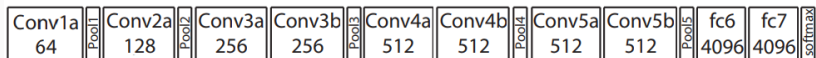
- Architettura jointFC → progettata ad-hoc per il task



- Pose di giunti come input
- Livelli completamente connessi (o densi)
- 2 livelli di Dropout
- Funzione di Softmax in output

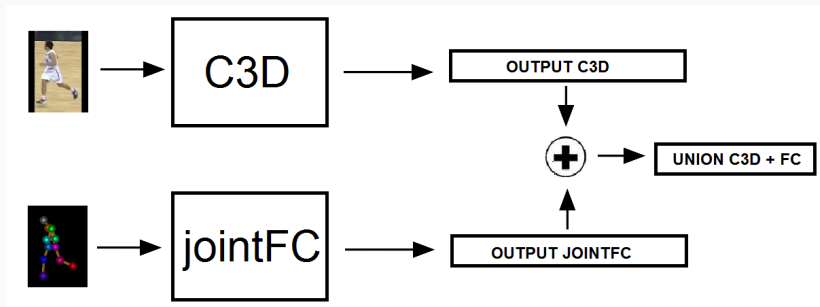
	Stoppata	Passaggio	Corsa	Palleggio	Tiro	Palla in Mano	Difesa	Blocco	Nessuna Azione	Camminata
Stoppata	0.54	0.04	0.10	0.05	0.08	0.06	0.06	0.00	0.01	0.06
Passaggio	0.06	0.15	0.12	0.19	0.05	0.29	0.03	0.01	0.01	0.10
Corsa	0.01	0.01	0.69	0.05	0.00	0.01	0.03	0.00	0.01	0.18
Palleggio	0.01	0.00	0.21	0.46	0.00	0.05	0.07	0.01	0.02	0.15
Tiro	0.11	0.07	0.09	0.04	0.48	0.15	0.02	0.00	0.00	0.04
Palla in Mano	0.05	0.03	0.05	0.06	0.04	0.56	0.07	0.01	0.04	0.08
Difesa	0.01	0.01	0.05	0.04	0.00	0.03	0.59	0.01	0.06	0.18
Blocco	0.05	0.01	0.05	0.14	0.01	0.12	0.10	0.18	0.14	0.18
Nessuna Azione	0.01	0.00	0.02	0.01	0.00	0.03	0.10	0.00	0.43	0.40
Camminata	0.00	0.00	0.11	0.02	0.00	0.02	0.06	0.00	0.08	0.70
Media	0.58									

- Architettura Convolutional 3D → Letteratura DNN



- Clips come input
- Kernel tridimensionali → features spazio temporali
- Livelli convolutivi + livelli completamente connessi
- Funzione di Softmax in output

	Stopata	Passaggio	Corsa	Palleggio	Tiro	Palla in Mano	Difesa	Blocco	Nessuna Azione	Camminata
Stopata	0.44	0.07	0.07	0.01	0.05	0.05	0.06	0.00	0.01	0.25
Passaggio	0.01	0.42	0.08	0.15	0.01	0.18	0.01	0.02	0.00	0.11
Corsa	0.00	0.01	0.87	0.02	0.00	0.01	0.01	0.00	0.00	0.09
Palleggio	0.00	0.02	0.04	0.89	0.00	0.02	0.01	0.00	0.00	0.02
Tiro	0.07	0.09	0.07	0.02	0.52	0.11	0.00	0.00	0.02	0.11
Palla in Mano	0.01	0.08	0.06	0.11	0.02	0.64	0.01	0.02	0.02	0.04
Difesa	0.00	0.01	0.04	0.02	0.00	0.03	0.53	0.02	0.11	0.24
Blocco	0.00	0.01	0.03	0.02	0.00	0.07	0.06	0.33	0.08	0.40
Nessuna Azione	0.00	0.01	0.00	0.00	0.00	0.01	0.04	0.01	0.66	0.28
Camminata	0.00	0.01	0.07	0.00	0.00	0.00	0.02	0.01	0.07	0.82
Media	0.73									



- Clip e posa di giunti in input.
- Classificazione mediante due componenti differenti:
jointFC + C3D

	Stopzata	Passaggio	Corsa	Palleggio	Tiro	Palla in Mano	Difesa	Blocco	Nessuna Azione	Camminata
Stopzata	0.44	0.05	0.07	0.03	0.05	0.05	0.08	0.00	0.01	0.22
Passaggio	0.00	0.36	0.07	0.18	0.01	0.23	0.01	0.01	0.00	0.15
Corsa	0.00	0.00	0.86	0.02	0.00	0.00	0.01	0.00	0.00	0.10
Palleggio	0.00	0.01	0.04	0.90	0.00	0.02	0.01	0.00	0.00	0.02
Tiro	0.11	0.07	0.11	0.02	0.48	0.11	0.00	0.00	0.02	0.09
Palla in Mano	0.01	0.06	0.07	0.12	0.01	0.64	0.02	0.01	0.02	0.04
Difesa	0.00	0.01	0.05	0.01	0.00	0.02	0.59	0.01	0.09	0.22
Blocco	0.01	0.00	0.03	0.03	0.00	0.06	0.08	0.31	0.09	0.39
Nessuna Azione	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.01	0.64	0.30
Camminata	0.00	0.00	0.07	0.00	0.00	0.00	0.02	0.01	0.06	0.83
Media	0.74									

Conclusioni

- 2 Dataset di Azioni Cestistiche di diverse natura
- Problema di sbilanciamento di dataset, con bassa accuracy con scarso numero di esempi.
- Confusione tra azioni simili tra loro (es. Palla in mano con Passaggio)
- Buone prestazioni di accuracy totale (74%), confrontata con lo stato dell'Arte per l'Action Recognition (85%)

Sviluppi Futuri

- Bilanciamento del Training delle Reti mediante Focal Loss
- Bilanciamento dei Dataset
 1. Modelli generativi per la creazione di nuovi esempi (GANs, variational autoencoders, ...)
 2. Nuove annotazioni manuali di azioni

Cliccare per il video

Grazie per l'attenzione