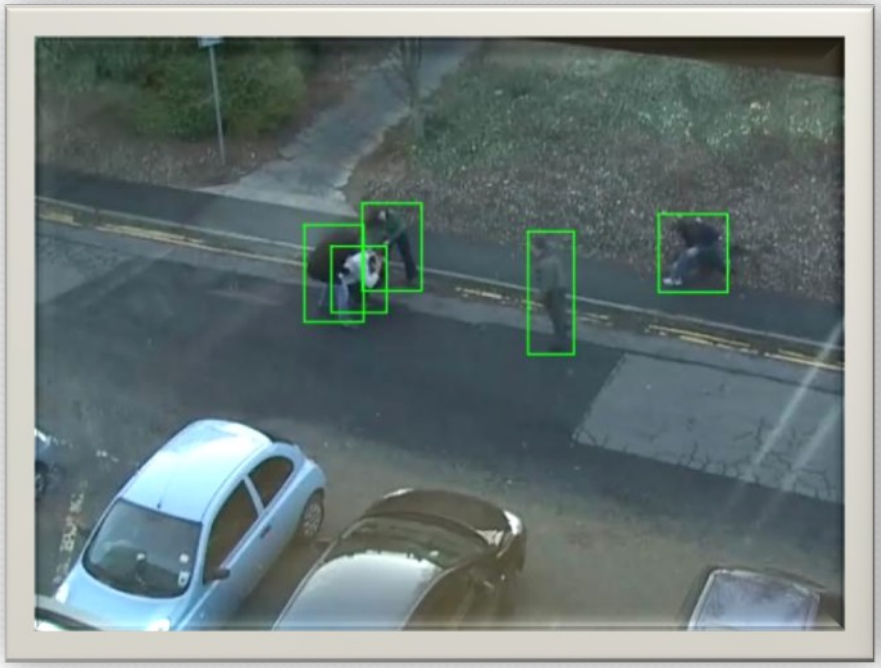
Classification of Pedestrian Interactions

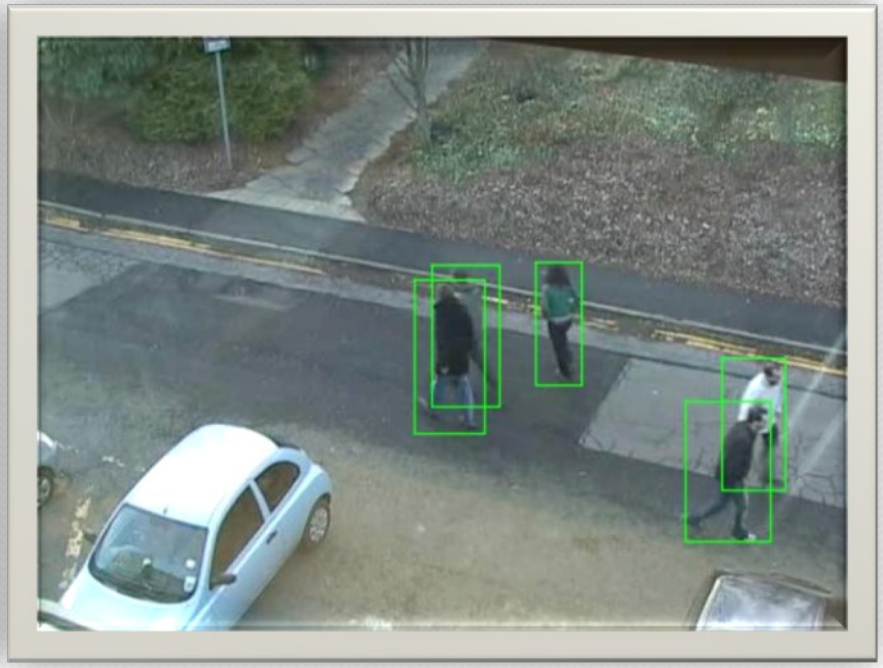
Scott Blunsden (me)
Bob Fisher
University of Edinburgh

What is this talk about?

- Considerations for metrics
- Present results, show considerations.
- Evaluations should give a good idea of the expected performance.

Main focus will be around the Edinburgh Dataset (see Bob's talk)





Plan

- Show you how we do classification.
- Then discuss issues surrounding how to evaluate it.

Classification

- Have sequences which are labelled.
  - Sequence labelled as a fight or people walking together.
- Divide sequences up into a training and testing set 50-50 split.

(Show Video)

Classification (2)

- Assume that
  - Tracking can be done reasonably well
- Features which are calculated
  - Speed of an individual time t
  - Alignment of two people (dot product)
  - Distance between people
  - Change in distance between people at time t and t-w
  - Difference in speed
  - Difference between the difference in position at time t and time t-w (are they getting nearer or further away)
  - Difference between starting positions (for an observed amount of time)
- We use PCA to reduce the dimensionality

Now for the results part

- I can get my method to classify

  97.93 % correct

- But what should I expect if I actually run this method ?
- And how useful is this statistic anyway?

(Show video)

What we really would like to know

- How well is this method doing
- How can we tell
- What are the main variables

Focus on the data from the video information here, not the model (although it is related).

What aspects of the data make the most difference?

- Ontology
  - What things are called how they are defined.
  - What is a behaviour ?
    - How do we define a fight? a meeting?
    - Here it is done by example (eg the labelled sequences define what we mean).
- Vocabularies may differ depending upon the user.
- What happened in the video depends upon what you were looking for.

Check Assumptions (make them available)

The Data itself

- No pre-defined test/training set? Should show error bars over multi runs.
  - Eg we took the best result, fine but what performance should I expect if I were to repeat this experiment.
- Agreed test set.
  - However you really want to know how well a method can be expected to perform.
- What's the expected performance?
- Confusion matrix and priors.
  - Per class performance is important. Frequent classes may dominate.

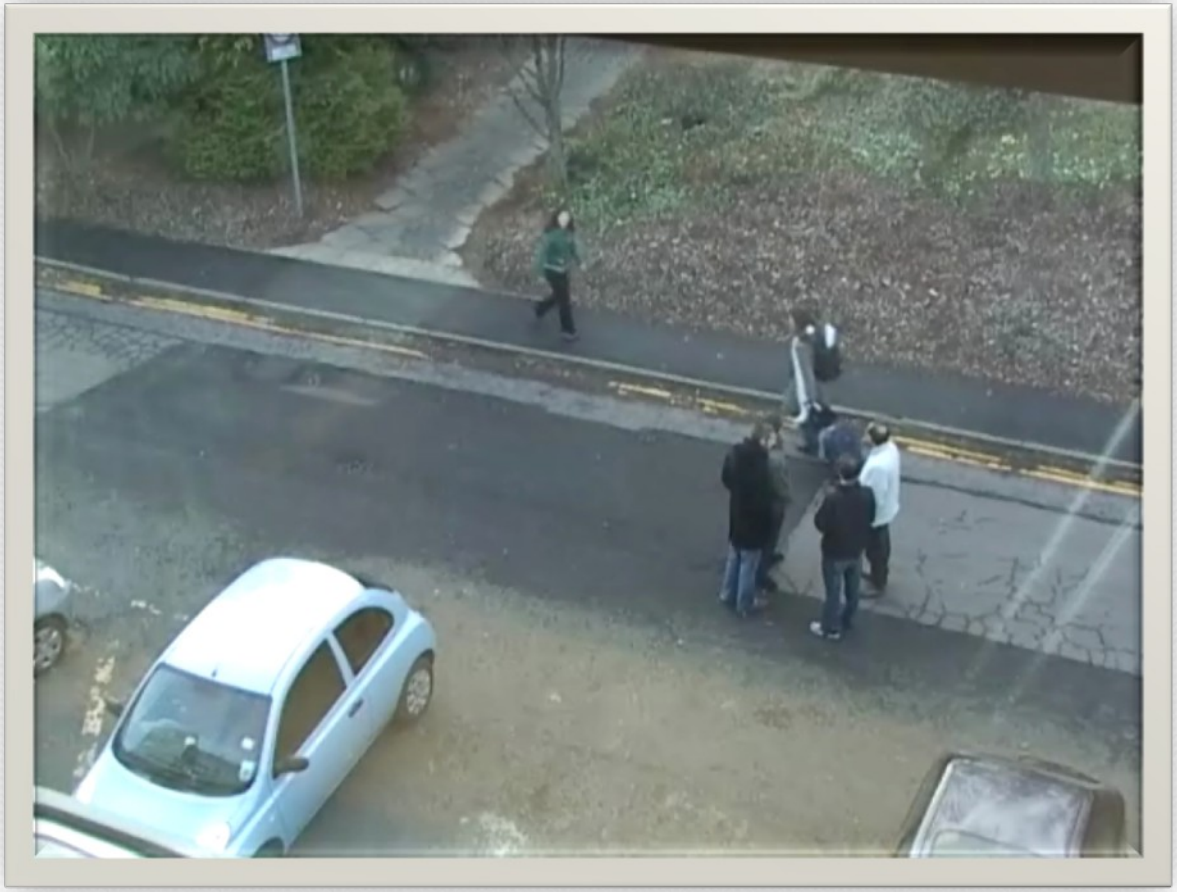
Some results

- Classification using a Conditional Random Field.
- Classify the pre-labelled sequences from the dataset.
  - Results are per frame:
    - mean: 96.03
    - min: 93.8
    - max: 97.93
    - var: 2.79

Confusion Matrix

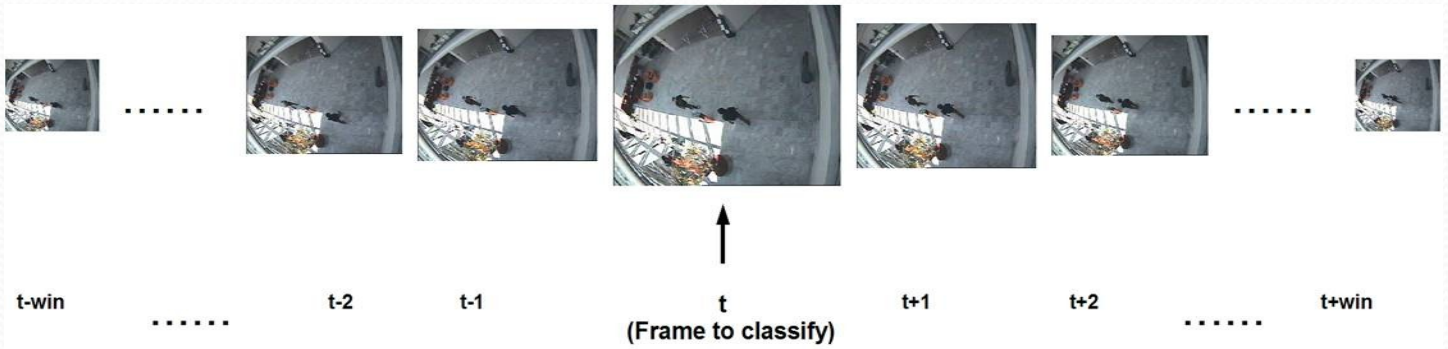
	In Group	Approach	Walk Together	Split	Ignore	Fight	Run Together	Chase
	43197	0	0	33	0	172	0	0
	68	3450	0	140	0	0	0	0
	71	0	5030	286	0	313	0	0
	0	0	0	2870	0	180	0	0
	0	236	10	0	289	0	0	0
	177	0	0	176	0	1514	0	0
	0	0	0	0	0	0	422	11
	0	0	0	0	0	0	65	35

Time

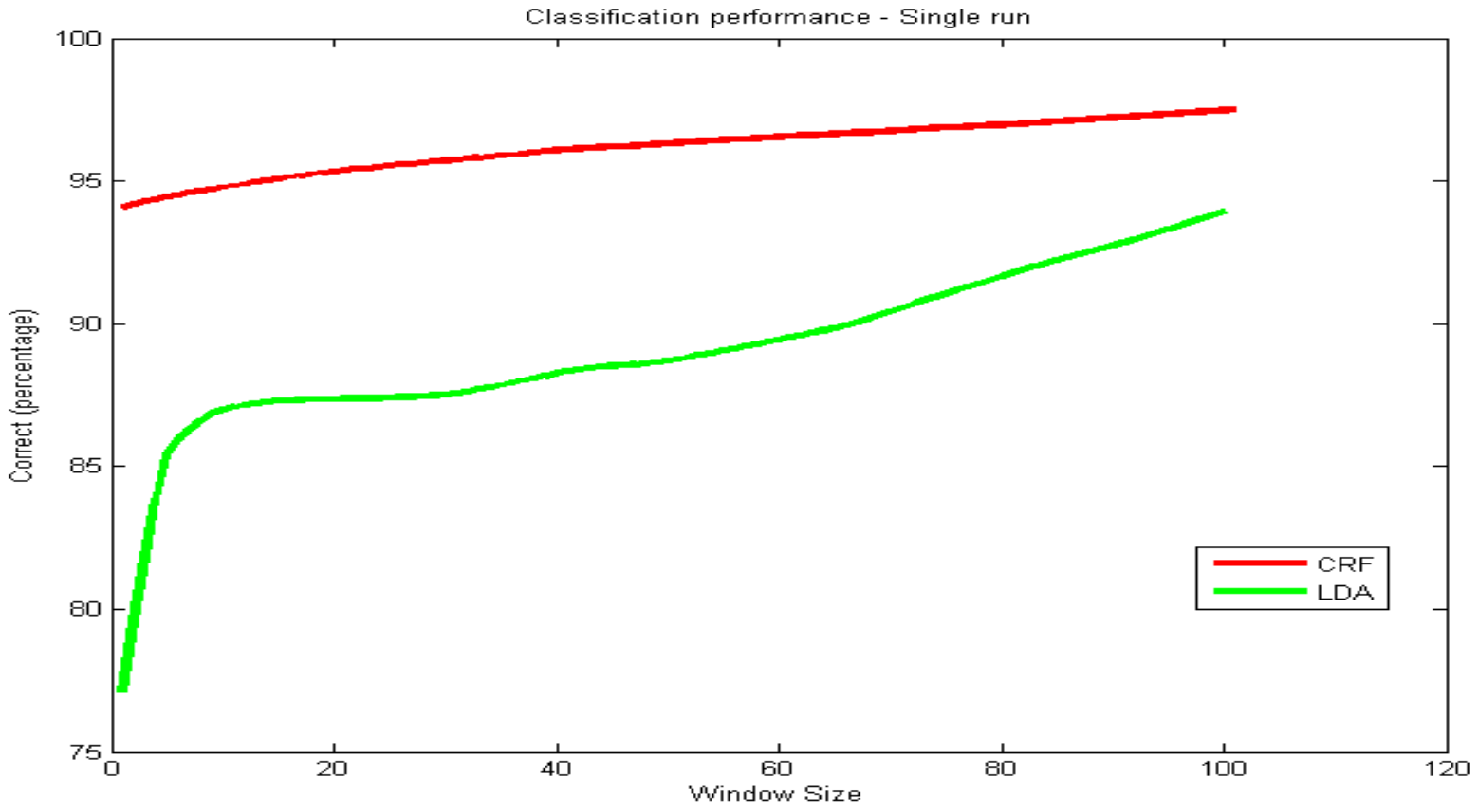


Time (2)

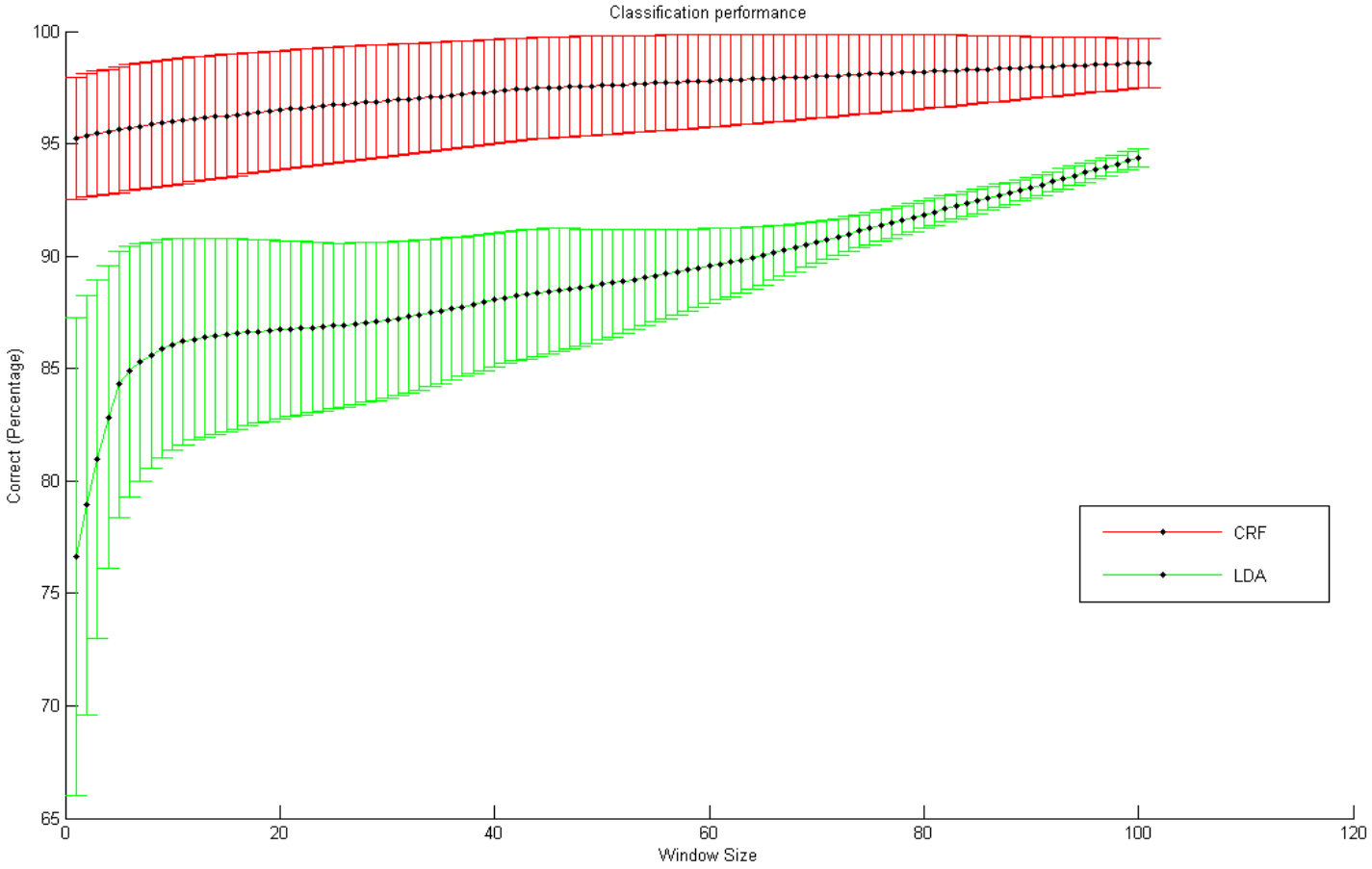
- There is a bound
  - Upper limit on how good accuracy can be given the amount of time you watch a sequence.
- There is a point where you can do no better.
- The longer the sequence the less examples you have for training.



Time (3) – Results - Single Run



Time (4) – Results – Multi Run



Things to consider

- Accessibility
  - Open and accessible data and labelling (others can check your assumptions).
- Consideration to training and testing sets.
  - Expected performance (rather than best).
- Bounds on information available (hard to determine but can be reported).