

# Can we date an artist's work from catalogue photographs?

## **Alex Brown**

Masters student, Computer Science,  
Aberystwyth University

## **Lloyd Roderick**

PhD student, School of Art, Aberystwyth  
University

and the National Library of Wales

## **Dr Hannah Dee**

Computer Science, Aberystwyth  
University

and **Professor Lorna Hughes**

The National Library of Wales



Sunset, Anglesey Sir John “Kyffin”  
Williams, 2004

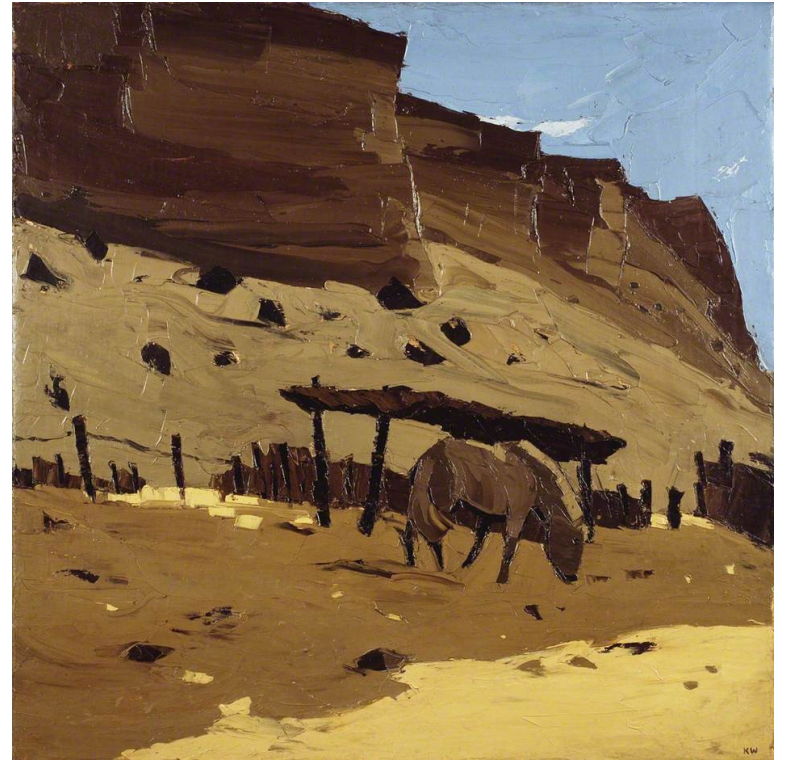




Snowdon, the Traeth, and the Frightened Horse. Sir John "Kyffin" Williams, 1948

# The man who painted in Welsh

- Sir John “Kyffin” Williams (1918-2006)
- Active for 6+ decades, with 2-3 paintings each week
  - Many paintings are in public collections
  - An unknown number are in private homes
- Epileptic
- Distinct **periods** in the work e.g. a spell in Patagonia (where it rains less than in Wales)



Lle Cul, Patagonia. Sir John “Kyffin” Williams, 1969



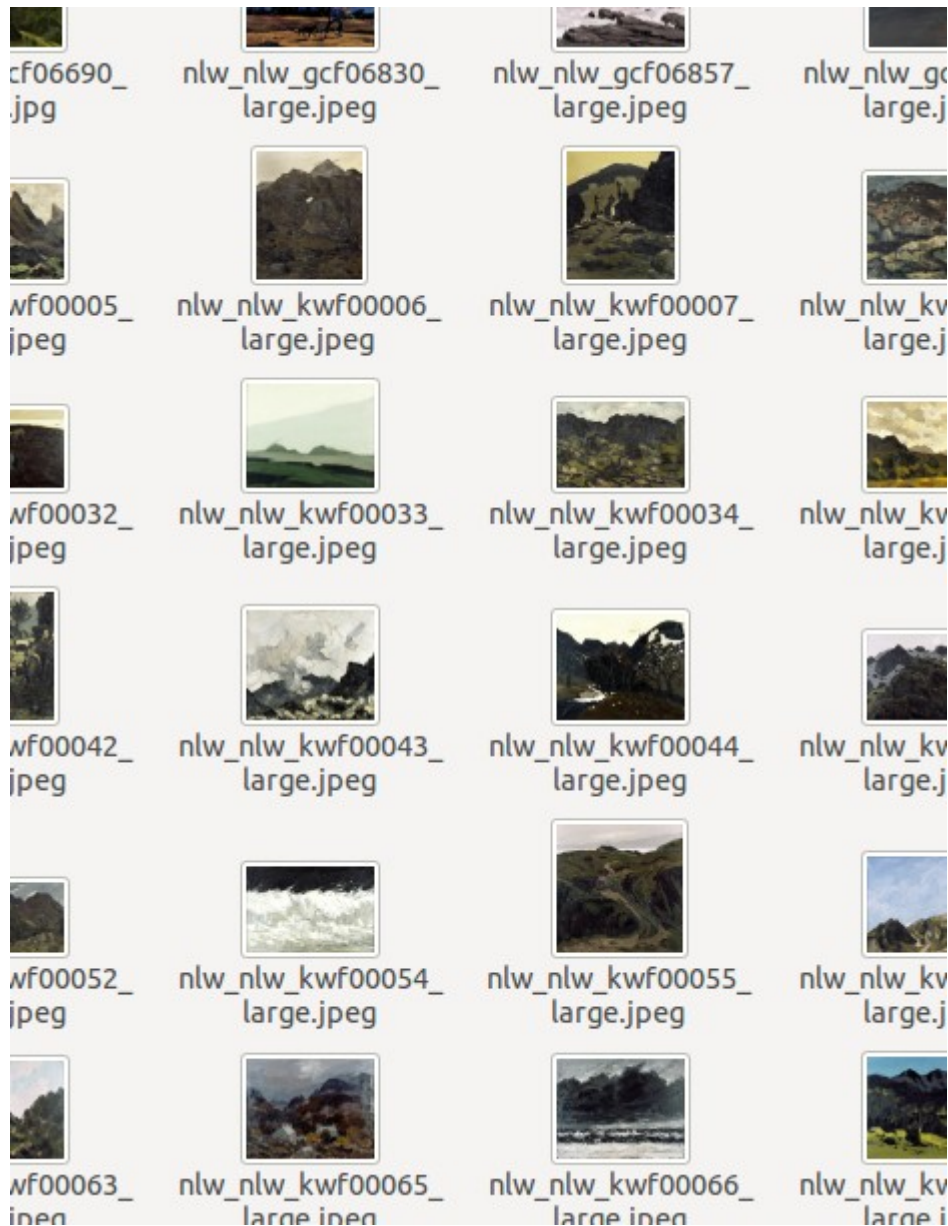
# Is it possible to guess the date of unknown paintings?

- For some paintings in public collections, dates are not known.
- Experts date this painting somewhere between 1970 and 1990
- Can we do better through computer vision?



Traeth Coch. Sir John "Kyffin" Williams, ????

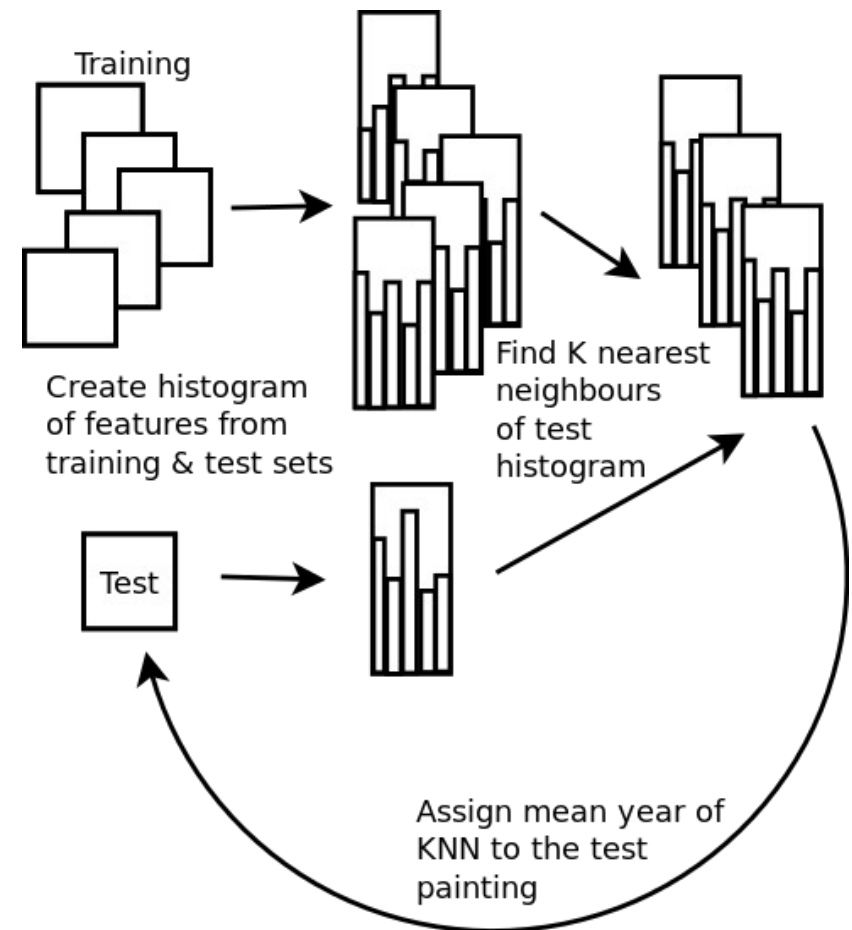
# Dataset facts and figures



- Over 300 Kyffin Williams artworks
- Collected online from catalogues and websites
- Various sizes & resolutions
- Uncontrolled capture environments
- We only use paintings where we know the date, which limits us to **102** paintings

# Leave-one-out Methodology

- Take a painting
- Work out statistical image representations for that and all the other paintings
- Find the nearest N other paintings in feature space
  - Using  $\chi^2$  distance
- Assign the mean year of these N paintings to the test painting



# Methodological considerations

- Allows us to maximise the power of the dataset
- Would be better if we had more to work with (more paintings, more known dates)
- Some popular years skew results
  - e.g. lots of paintings in 1969/1970
- We have actual and predicted year for all paintings
- Thus we can correlate these (using Pearsons  $r$ ) and test for significance

# Image processing: colour

- Kyffin Williams' paintings are often described as **gloomy**
- Does this change with time?
- We used colour-based features to investigate this
- Despite uncontrolled image capture settings, we found positive correlations with all three measures

## COLOUR features

Mean RGB

Mean HSV

Colour histograms



# Image processing: lines and edges

- Kyffin Williams' paintings seem to get blockier
- In order to try and use this observation we investigated two main line-based features
- These capture a rough mixture of **how many** edge pixels there are in a painting, and **what orientation** these are in

## EDGE features

Count of Canny edge pixels

Steerable filters

# Image processing: texture descriptors

- We used Gabor filters, and discrete derivative filters
- Both of these were used to build a histogram of line orientations representing the dominant angles of paint strokes
- In this work Gabor bandwidth was fixed, and orientation varied

## **TEXTURE features**

Gabor filters

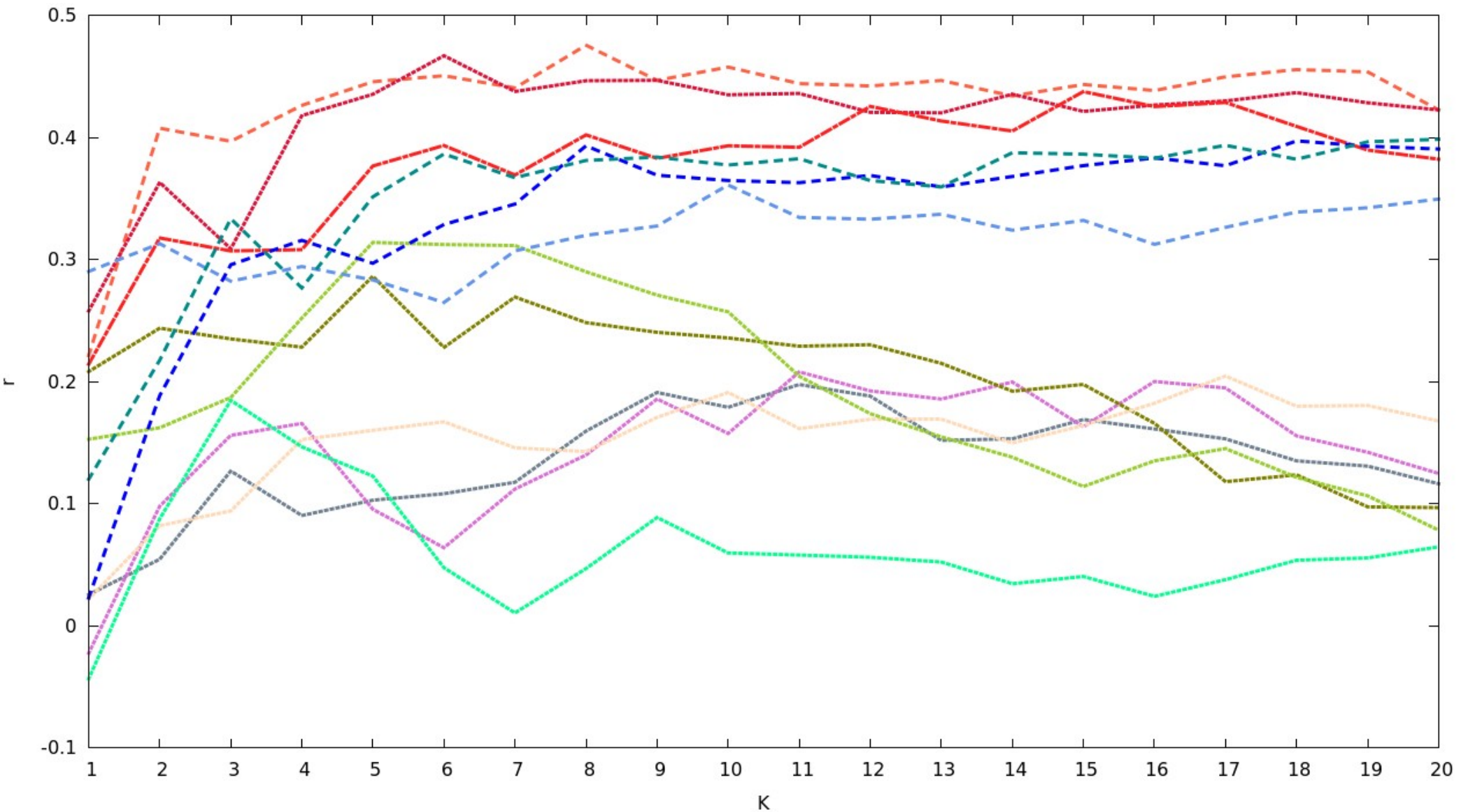
Discrete  
derivative filters  
(used to build a  
histogram of  
oriented  
gradients)

# Evaluation

- There is no “gold standard” to compare against
- We use a leave-one-out methodology, combined with K-nearest neighbours to assign dates.
- Thus K is a parameter we need to consider
  - $K = 1$  has the effect of assigning the year of the nearest painting in feature space
  - $K = 102$  has the effect of giving each painting the date of the mean year
- By graphing the correlation of predicted year against actual year, we can investigate what values of K work best here

# A graph of K against r

RGB  
HSV  
RGB Histograms  
HSV Histograms  
Edge Strength  
Steerable filters: 4 orientations  
Gabor filters: 4 orientations  
Gabor filters: 8 orientations  
Gabor filters: 16 orientations  
HOG (Discrete Derivatives): 4 orientations  
HOG (Discrete Derivatives): 8 orientations  
HOG (Discrete Derivatives): 16 orientations





# Another way to evaluate prediction

- It seems that a K value around 7 or 8 works best for most of our feature sets
- Using this value, and our best performing metric, we can come up with another evaluation method
  - Based upon difference between predicted and actual year
- $C(n)$  is the percentage of paintings for which our classifier manages to date the artwork in question within  $n$  years of actual painting date.
- Our best method (Gabor filters, 4 orientations) can date 71% of paintings to within 15 years.

# Expert knowledge

- Artistic exemplars: ask an expert (Dr Paul Joyner, of the National Library of Wales)
- Statistical exemplars: Closest painting to centroid in feature space for that year.
- Rather than use K-nearest neighbours, use closest exemplar

Method	r	P(r)	C(15)
Artistic exemplar	0.328	<0.0001	57%
Statistical exemplar	0.383	<0.0001	61%
Centroid	0.403	<0.0001	64%

# So...

Can we date  
an artist's  
work from  
catalogue  
photographs?

## Sort of.

- We get strong correlations between predicted and actual date
- Our guess falls within 15 years for the majority of paintings (71%)

# Issues and further work

- Dataset issues
  - Size (102 paintings)
  - Capture issues
- Impossibility of comparison
  - if you find a scholar who can date a painting, they know the work already
- Feature selection and combination
- Try a different classifier
  - Nearest Neighbour problematic with a sparse dataset
- Test on another artist
  - Simple features
  - No artist specific models
- Incorporation of other metadata e.g. size
  - Location is of interest





Sir John "Kyffin" Williams - Ceg y Ffos, 1969









# Summary:

A bag-of-visual-words approach to dating paintings, using KNN to guess the year

## Any questions?

Contact:

Hannah Dee

[hmd1@aber.ac.uk](mailto:hmd1@aber.ac.uk)



Image: Portrait by Bernard Mitchell,  
<http://www.llgc.org.uk/index.php?id=httpwwwllgcorguktypo3sys>