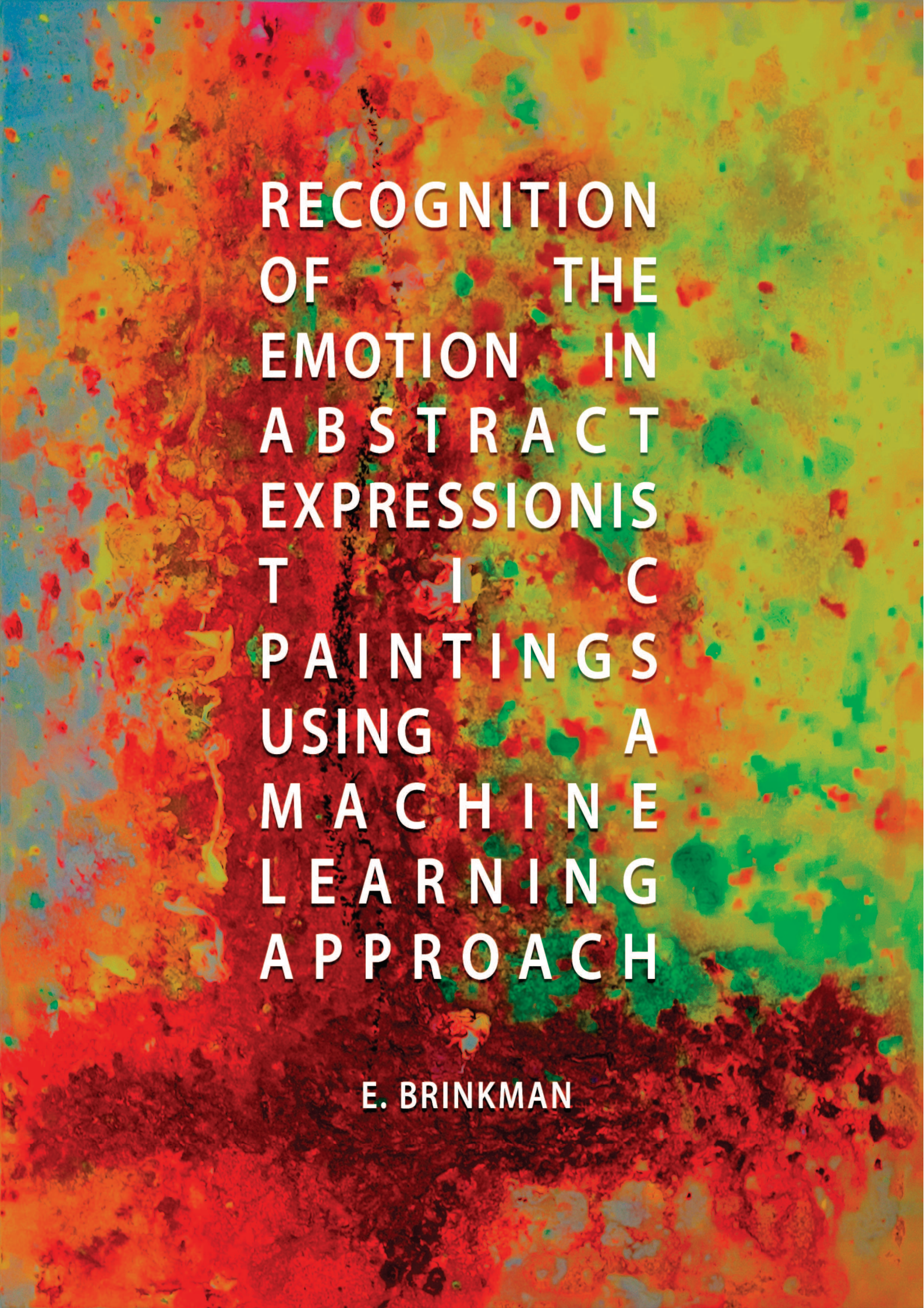

RECOGNITION OF THE EMOTION IN ABSTRACT EXPRESSIONISTIC PAINTINGS USING A MACHINE LEARNING APPROACH
E. BRINKMAN

MSC THESIS

RECOGNITION OF THE EMOTION IN ABSTRACT EXPRESSIONISTIC PAINTINGS USING A MACHINE LEARNING APPROACH

E. Brinkman

Department of Human Media Interaction, Twente University, Enschede, The Netherlands

Supervisors: Z. Ruttkay, M. Poel, D. Heylen

Date: January 18, 2008

ABSTRACT

The topic of the research described in this master thesis is: "A machine learning approach to emotion classification of abstract expressionistic paintings." Emotion in paintings is a result of the sum off all the elements within a painting. In abstract expressionist paintings abstract shapes, expressive colors and textures depict emotion in the painting. The research involves the abstract expressionistic art movement, evaluation of human emotions and machine learning algorithms.

The main points of interest within this research is the relation between abstract expressionistic paintings and the emotion they are assumed to evoke at human observers. This leads to the following research question: "Is it possible to predict how people would react emotionally when observing an abstract expressionistic painting?"

The steps followed to conclude on this research are as following. An emotion corpus on how human observers react emotionally when observing abstract expressionistic paintings is collected and analyzed. This emotion corpus contains ratings on three emotion dimensions; arousal, dominance and valence. Image features on a set of abstract expressionist paintings are extracted and analyzed. Next the relation between the image features and the emotion corpus is analyzed. Machine learning algorithms are used to evaluate their ability to classify abstract expressionistic paintings on how people would react emotionally when observing these paintings.

The results show that prediction of the evoked emotion by classification algorithms give a correct classification of 73.33% (MAE 0.16) on the arousal dimension, 46.67% (MAE 0.28) on the dominance dimension and 73.33% (MAE 0.28) on the valence dimension. It shows that the arousal dimension the easiest emotion dimension to predict an evoked emotion on. These results are found after applying feature selecting to reduce the number of image features used to train the classifier. This results in a number of image features that are most valuable when predicting the emotion of an abstract expressionistic painting.

On the arousal emotion dimension, edges found in the paintings' images are believed to be a fair predictor. Also diversity within the intensity of a color does seem to have a strong relation to emotion and color by itself does not. This is in contradiction as what is commonly believed, e.g. to see red as an angry color and yellow as a happy color.

The dominance emotion dimension does not show good results when predicting with the classifiers. Also most of its related image features are also used in relation to the arousal emotion dimension. It is questionable if the dominance dimension is a good measure for emotion in abstract expressionistic paintings.

The predicting feature on the valence emotion dimension needs more ground to be able to conclude on its relation. When accepting the found relation it would mean that absence of very small horizontal lines leads to a painting to be less pleasant to observe.

PREFACE

Dit werk is geschreven als onderdeel van de masteropleiding "Human Media Interaction". Deze opleiding wordt door mij gevolgd aan de Universiteit Twente te Enschede. Mijn interesse voor het onderwerp dat hier wordt belicht komt voort uit mijn ervaring met informatica en automatisering op een professioneel gebied. Ook de overvloedige aanwezigheid van allerlei media in mijn omgeving heeft bijgedragen aan mijn beslissing om mij binnen dit gebied te verdiepen. Waar bij velen de interesse voor een bepaald gebied voortkomt uit waardering is dit in mijn geval niet zo. Persoonlijk had ik vooral een afkeer van nieuwe media en de daaraan gerelateerde consumptie en gedragspatronen. Het is dan ook vooral het onbegrip en het niet begrijpen van het waarom wat mij ertoe heeft aangezet mij te willen verdiepen in dit gebied.

Een ander vlak wat mij heeft bewogen om dit werk te maken is mijn waardering voor artiesten en de werken die door deze gemaakt zijn. De overtuiging van een artiest is vaak in een kunstwerk terug te vinden. Deze eerlijkheid en verdieping herbergt een grote schoonheid. Op het vlak van de abstract expressionisme hebben de werken naar mijn persoonlijke mening een zware emotionele uitwerking. Naar aanleiding van deze ervaringen is het voor mij interessant om dit werk te schrijven.

Een laatste motivatie voor het onderzoeken van de relatie tussen emoties en beeld is de aandacht die door de opleiding is gegeven aan het menselijke aspect binnen mens en media interactie. De emoties van de mens spelen hierbij een grote rol, deze rol heeft dan ook terecht veel aandacht gekregen binnen het vakgebied.

Van mijn kant wil ik dan ook de opleiding bedanken voor een zeer actuele visie op de huidige ontwikkelingen binnen het vakgebied. Ook de kunde van de medewerkers binnen de opleiding is opvallend van een hoog niveau. Mijn begeleiders wil ik graag bedanken voor hun geduld en hulp tijdens het proces van totstandkoming van dit werk.

Ook wil ik mijn klasgenoten bedanken voor de prettige werksfeer tijdens de verschillende vakken en projecten. Verder nog mijn vrienden voor het geven afleiding, inspiratie en ontspanning en mijn familie voor de zorgzaamheid die zij geven.

TABLE OF CONTENTS

Abstract 2
Preface 3
Table of contents 4
List of figures 5
List of tables 5
1. Introduction 6
1.1. Background 6
1.1.1. Abstract expressionism 6
1.1.2. Image features of paintings 8
1.1.3. Emotion recognition and stimuli classification related 9
1.2. Structure of the report 11
2. Methodology 12
3. Data acquisition 14
4. Feature extraction 15
5. Data evaluation and cleaning 18
6. Machine learning 19
7. Test results 21
7.1. Diversity of paintings 21
7.2. Evaluation of the emotion corpus 21
7.3. Correlation analysis 26
7.4. Comparison of machine learning algorithms 30
8. An example on the emotion classification of websites 33
9. Conclusions and recommedations 34
References 36
Appendices 38

LIST OF FIGURES

Figure 1 - distribution of ratings 22

LIST OF TABLES

Table 1 - emotion dimension, rating scales and their meaning 14
Table 2 - kappa values on 5 categories 21
Table 3 - agreement and distribution on the arousal scale 22
Table 4 - agreement and distribution on the valence scale 23
Table 5 - agreement and distribution on the dominance scale 23
Table 6 - words related to emotion dimensions 24
Table 7 - P-values on correlation between the mode and statistics of target data 25
Table 8 - P-values on correlation between the mean and statistics of target data 25
Table 9 - P-values on correlation between the standard deviation and statistics of target data 25
Table 10 - P-values on the correlation between entropy values of colors and emotion dimensions 26
Table 11 - P-values on the correlation between mean values of colors and emotion dimensions 27
Table 12 - P-values on the correlation between standard deviation of colors and emotion dimensions 27
Table 13 - numbers of correlating features 29
Table 14 - classifier efficiency 32
Table 15 - accuracy and MAE on human raters 32
Table 16 - accuracy and MAE of emotion classification of websites 33
Table 17 - confusion matrices of emotion classification of websites 33

1. INTRODUCTION

Emotion in paintings is a result of the sum off all the elements within a painting. Classic paintings often depict human beings where facial expressions, gestures and scenes evoke most of the emotion at the observer. In abstract expressionist paintings abstract shapes, expressive colors and textures depict emotion in the painting. The placing of the shapes and selection of the colors and textures seem random but are as often carefully arranged by the artist.

The topic of the research described in this document is: "A machine learning approach to emotion classification of abstract expressionistic paintings." This can be explained as using computer algorithms to automatically judge an abstract expressionistic painting on the emotion as it would be evoked at human observers. The word automatically applies the use of computer algorithms that are in some extend able to learn, hence a machine learning approach to emotion classification. The research involves the abstract expressionistic art movement, evaluation of human emotions and machine learning algorithms.

1.1. BACKGROUND

This section contains work that is related with abstract expressionism, image features of paintings, evaluation of emotions and automatic classification of images.

1.1.1. ABSTRACT EXPRESSIONISM

The history of abstract expressionism starts around 1910. The painter Wasily Kandinsky was one of the first western painters who made non-figurative and non-objective paintings. He was also the first to use the term 'abstract expressionism' in 1918. In his work the emotion within the artwork and the artist's inner need are central subjects. For Kandinsky color was his language and the most important voice of the painting as tone is for music. Some quotes by Kandinsky give a good impression of his vision.

> *"I applied streaks and blobs of colors onto the canvas with a palette knife, and I made them sing with all the intensity I could..."*
>
> *"Color provokes a psychic vibration. Color hides a power still unknown but real, which acts on every part of the human body."*

Also Kandinsky relates color to emotions, giving meaning to the color, as the following quotes tell.

> *"Almost without exception, blue refers to the domain of abstraction and immateriality."*
>
> *"As a picture painted in yellow always radiates spiritual warmth, or as one in blue has apparently a cooling effect, so green is only boring."*
>
> *"Absolute green is the most restful color, lacking any undertone of joy, grief, or passion. On exhausted men this restfulness has a beneficial effect, but after a time it becomes tedious."*

For Kandinsky it was clear that color can not only be used to express an inner need, but also to evoke an esthetical reaction and maybe possible a mental reaction to the beholder of a painting.

Due to the World War II many European artists fled to the United States of America. Among them were surrealistic painters who's ideas influenced a young generation of painters who were struggling to find a voice for American art. This influence caused a new movement, later known as abstract expressionism. Important names within this movement are Marc Rothko, Willem de Kooning and Jackson Pollock. The movement was exceptionally successful and still an influence for many artists today.

The philosophy of abstract expressionism has been discussed many times within the movement. As the following quote from Hans Hoffman and Willem de Kooning on a three-day closed conference shows, dated April 20-23, 1950.

> De Kooning*: "I wonder about the subject matter of the Crucifixion scene – was the crucifixion the subject matter or not? What is the subject matter? Is it an interior subject matter?"*
>
> Hoffman: *"I think the question goes all the time back to subject matter. Every subject matter depends on how to use meaning. You can use it in a lyrical or dramatic manner. It depends on the personality of the artist. Everyone is clear about himself, as to where he belongs, and in which way he can give aesthetic enjoyment. Painting is aesthetic enjoyment. I want to be a 'poet'. As an artist I must conform to my nature. My nature has a lyrical as well as a dramatic disposition. Not one day is the same. One day I feel wonderful to work and I feel an expression which shows the work. Only with a very clear mind and on a clear day I can paint without interruptions and without food because my disposition is like that. My work should reflect my moods and the great enjoyment which I had when I did the work."*

Color field and gestural painters are both seen as abstract expressionistic painters. Both share an identical philosophy, but their approach differs. The following quote is from Meyer Schapiro, an art historian, dated 1956.

> *"Each seeks an absolute in which a receptive viewer can lose himself, the one in compulsive movement, the other in an all-pervading, as internalized, sensation of dominant color. … The result in both is a painted world with a powerful immediate impact."*

On the approach of gestural painters, a quote from Jackson Pollock.

> *"when I am in my painting. I'm not aware of what I'm doing. It is only after a sort of 'get acquainted' period that I see what I have been about. I have no fears about making changes, destroying the image, etc., because the painting has a life of its own. I try to let it come through."*

To illustrate the approach of color field painters, some quotes from Marc Rothko.

> *"I am not an abstract painter. I am not interested in the relationship between form and color. The only thing I care about is the expression of man's basic emotions: tragedy, ecstasy, destiny."*
>
> *"only in expressing basic human emotions — tragedy, ecstasy, doom, and so on. And the fact that a lot of people break down and cry when confronted with my pictures shows that I can communicate those basic human emotions . . . The people who weep before my pictures are having the same religious experience I had when I painted them. And if you, as you say, are moved only by their color relationship, then you miss the point."*

This short introduction on abstract expressionism shows that different artists within its history all seek to express abstractions as emotions and the so called inner need. Some

use a more conscious approach as Kandinsky and Rothko, others as Pollock use a more unconscious approach. This research can be seen as a more scientific reflection on the art movement and their urge to express themselves.

1.1.2. IMAGE FEATURES OF PAINTINGS

This section discusses image features of paintings as they are used by painters, not the features that can automatically be extracted from an image. It can be seen as an introduction of the visual language a painter uses in a painting.

The basic features of paintings that are of interest for this research are color, texture, shape, composition and contrast. Features that are related to figurative paintings as narrative content and symbolism are not discussed here. These are outside of the scope of the research. Although figures like a head or a building can be seen in some abstract expressionistic paintings, it is most likely not the painters intention to paint these figures.

Color - Painters often use a color model named RYB, an abbreviation of red yellow and blue. These colors are subtractive primary colors. By mixing these primary colors the secondary colors violet orange and green are made. By mixing a secondary color with a primary color the tertiary colors are made, examples are orange-red, red-violet and blue-green. Mixing a color with a complementary color gives a shade of gray. Brighter colors are made by mixing a color with white. Darker colors can be made by mixing a color with black or a complementary color.

The meaning of a color depends on it's context and the culture of the context. For example, color preferences of Korean people differ from the preferences of English people. Korean people prefer soft , bright and pastel tone colors, whereas English people prefer strong, dark, dull and intense colors [CHOI].

Texture - Texture in the art of painting is the feel of the canvas. Texture is based on the paint and its application or addition of materials such as paper, metal, wood, silicon, sand, etc. Texture in a painting stimulates two different senses, vision and touch. This makes it a unique element of art. Ocvirk et.al. describe 4 types of texture in art: actual texture, simulated texture, abstract texture, and invented texture [OCVIRK].

- Actual Texture, the actual texture of how the work of art looks and feels to the touch.
- Simulated Texture, creating a visual effect of texture without actually adding texture.
- Abstract Texture, texture that does not seem to match the object its connected with.
- Invented Texture, the creative way of adding alternate materials to create an interesting texture.

Shape - A shape is the appearance of something defined by its outline. For example, a yellow area on a red canvas creates two shapes. The outline is found between the yellow and red areas.

Composition - Composition is the arrangements of elements in an artwork, it defines the position and size of shapes and colors.

Contrasts - Artists use all kinds of contrasts to achieve certain effects. Best known are: color-against-color contrast, light-dark contrast, cold-warm contrast, complementary contrast, simultaneous contrast, quality contrast and quantity contrast. Basically, there is a contrast known for almost any value that can be given to color, shape and

composition. The effects of these theories tend to differ between cultures and fashions [ITTEN].

Jocobson and Bender discuss that although it is possible that one viewer's perception of color may be very different from another's, the relationships between colors are, in many respects, universal. They describe a model of color experience that is based on types of interactions among colors. It adjusts attributes as hue, value, saturation and their contrasts, as well as size and proportion [JACOBSON].

1.1.3. RELATED WORK ON EMOTION RECOGNITION AND STIMULI CLASSIFICATION

This section discusses work related to this research that uses computer algorithms to predict the emotional value or classify other features of a stimuli. These works are found in the field of speech and language and image processing.

On emotion models

Emotion models are developed to represent an emotion or a feeling of humans, not how emotion and feeling arise.

Basic emotions - In general the models that use words to differentiate emotion categories were traditionally used for psychological studies. For example; Ekman describes basic or universal emotions; anger, fear, sadness, happiness and disgust. These are emotions that can be found within every human and are believed to be primary, innate, and are universal [EKMAN].

The advantage is that basic emotions are well known by the people. The disadvantage is that one can feel other emotions than those basic ones, which have to be mapped on the model's categories and leads to some distortion of the actual impression. Further, different researchers might use different basic emotion categories, which makes comparing studies difficult. Also, they are found to be problematic, especially when dealing with cultural differences and translations.

Emotion dimensions - Models that use a dimensional approach to represent the emotion are found to overcome cultural difficulties. This because translations of the words that describe the emotion categories can have a different meaning in other languages. By using the dimensions of arousal and valence the measurement of the felt emotion can be very accurate. The disadvantage is that the dimensions are labeled with rather abstract terms and requires understanding by the rater.

The dimensional model represents emotional states in terms of multiple dimensions, arousal and valence. Arousal measures how dynamic the emotional state is. For instance, exhilaration involves a very high level of arousal, boredom involves a very low one. Valence is a measure of the positive or negative feeling associated with the emotional state. For instance, happiness involves a very positive valence, despair involves a very negative one. Many techniques converge on the conclusion that to a first approximation, emotion terms can be understood as referring to points within a circular space defined by those two dimensions [SCHLOSBERG, PLUTCHIK, ROSEMAN, RUSSEL].

The two-dimensional model can be extended with other dimensions, an often used dimension is the control level (or the dominance-submission dimension) [COWIE]. It is said to make it possible to distinct fear from anger [KEHREIN].

On speech and language

Related is the research done on emotion recognition in non-acted speech by Steidl et. al.[STEIDL]. They use an entropy-based measure to find the unanimity of the human raters and machine classifiers. This because human raters do not agree on one common class in most cases. They show that in their case, a classifier which achieves a recognition rate of only 60% performs as well as five human raters on average.

Petrushin [PETRUSHIN] did research on agents for emotion recognition in speech and their application to a real world problem. The agents can recognize five emotional states with the following accuracy: normal or unemotional state - 55-75%, happiness - 60-70%, anger - 70-80%, sadness - 75-85%, and fear - 35-55%. The total average accuracy is about 70%.

Batliner et. al.[BATLINER] used realistic, spontaneous speech data from an AIBO corpus. The emotional annotations of five labelers, transformed into a confusion matrix, were used in a non-metrical multi-dimensional scaling to display two dimensions, the first being valence. The second, however, not arousal but interaction, i.e., addressing one self (angry, joyful) or the communication partner (motherese, reprimanding). They show that it depends on the specifity of the scenario and on the subjects conceptualizations whether this new dimension can be observed, and discuss impacts on the practice of labeling and processing emotional data.

In the context of a prototype voice portal Burkhardt et.al.[BURKHARDT] describe methods to search for training data, report on the performance of the classier under real world conditions and explore the use of dialog information for anger prediction. The results show that, although significantly worse than under laboratory conditions, anger detection still works well above chance level and can be used to enhance real world voice-portal usability.

On images

Art and paintings - Research on comparing art by computer classification is mostly focused on identification of the painter of an artwork [HERIK, KEREN, WIDJAJA, SABLATNIG] and authentication [KRONER, LYU] of an artwork. Some work is done on classification of paintings on art movements [ICOGLU] and style[KIRCH].

Sablatnig et. al. proposes a method that relates to certain characteristics of an art work to a specific artist. In order to study this personal style, they examine brush strokes in particular in portrait miniatures.

Image retrieval - The research field of content based image retrieval and image classification did a lot of work on feature extraction of images [SMEULDERS]. Most of the methods extract features on color, shape and texture information.

Some interesting feature extraction methods use the histogram of an image. Based on this approach the correlogram [HUANG], a three dimensional histogram that holds the spatial distance between color pairs and the geometric histogram [RAO] that combines the histogram and the correlogram are developed.

Emotion classification - According to the strong relationship between colors and human emotions, an emotional semantic query model based on image color semantic description is proposed by Wang and Yu [WANG04]. First images are segmented into regions through a color image segmentation algorithm. Then, term sets are generated through a fuzzy clustering algorithm so that colors can be interpreted in semantic terms. They extended the method to extract the color semantic of image regions, and developed an approach for describing image color semantic including regional and global semantic description.

Wang et. al. [WANG05] propose an approach to classify art paintings into emotional categories (dynamic vs. static). The key points are feature selection and classification algorithms. According to the strong relationship between notable lines of an image and human sensations, a novel feature vector WLDLV (weighted line direction-length vector) is proposed, which includes both orientation and length information of lines in an image.

On facial expressions

Bartlett et.al. present a systematic comparison of machine learning methods applied to the problem of fully automatic recognition of facial expressions. The facial expressions are coded into 7 dimensions in real time: neutral, anger, disgust, fear, joy, sadness and surprise. The best results were obtained by selecting a subset of Gabor filters using AdaBoost and than training support vector machines on the outputs of the selected filters. The generalization performance to new subjects for recognition of full facial expressions in a 7-way forced choice was 93% correct [BARTLETT].

1.2. STRUCTURE OF THE REPORT

The following chapters of the report contain the methodology of the research, evaluation of the results and conclusion and recommendations. The methodology states the research questions and explains the design of the research and decisions made along the way. The results contain evaluation of data gathered by conducting experiments. The conclusions and recommendations will discuss and conclude on the research topics, also recommendations for further research within this topic are given.

2. METHODOLOGY

The methodology of this research contains the research goals, a definition of the scope of the project, a number of research questions and the design of the research.

The main points of interest within this research is the relation between abstract expressionistic paintings and the emotion they possibly evoke at human observers. By using computer technologies this research aims at a machine learning approach that can be used to automatically classify these paintings on the emotion they evoke. This research can be of use in any domain where abstract non-figurative images are presented to human observers. Think about judging interiors of buildings, backgrounds of computer games and designs of websites.

The scope is the field of abstract expressionistic paintings. These paintings are used because of the link between the philosophy of the movement and human emotions; they are created because of the inner need of the painter to express his feelings within the artwork. Also, most of these paintings do not contain figurative images as facial expressions, human gestures or dramatic scenes. These figurative images do effect the emotion evoked on the human observer. By using the scope of abstract expressionistic paintings this research can focus on features as color, shape, contrast and composition and their relation to the emotion evoked.

The research topic with respect to the stated goals leads to the following research questions:

- Is it possible to predict how people would react emotionally when observing an abstract expressionistic painting?

The research question involves the following sub questions:

- How do human observers react emotionally when observing an abstract expressionistic painting?
- How are the features of abstract expressionistic paintings related to the evoked emotion at the human observers?

These research questions are made with one assumption: abstract expressionistic paintings do evoke an emotion at their observers. This is assumable because of the popularity of the art movement, its philosophy and reviews by art critics.

To be able to conclude on the research questions the following methodology is used. It describes a number of steps to follow and the way they are implemented. This includes possible alternative methods that can be used within these steps and assessments on alternatives. The methodology includes the following steps:

- Data acquisition
    - To collect an emotion corpus on how human observers react emotionally when observing abstract expressionistic paintings.
- Image feature extraction
    - To collect image features on a set of abstract expressionist paintings.
- Data evaluation and cleaning
    - To analyze the image features of abstract expressionist paintings.

    - To analyze the emotion corpus on how human observers react emotionally when observing abstract expressionistic paintings.
    - To analyze the relation between the image features and the corpus.
- Machine learning
    - To evaluate the ability of machine learning algorithms to classify abstract expressionistic paintings on how people would react emotionally when observing an these painting.
    - To evaluate the ability of machine learning algorithms to classify images outside the scope of abstract expressionistic paintings on how people would react emotionally when observing these images.

These steps are explained in the following chapters.

3. DATA ACQUISITION

Knowing the emotion evoked by paintings at human observers is most important for this research. When attempting to use a machine learning approach to conclude on the possibility to predict the evoked emotion, the evaluation data must have some agreement between the observers.

Also, a machine learning approach demands training data on the subjects, hence the paintings. Therefore this experiment will be used to gather an emotion corpus as training data; the evoked emotion as rated by the human observers.

A questionnaire is used to gather data on the evoked emotions. This form is chosen because it can be implemented online and thus gives the opportunity to reach a large amount of people. Also, it can generate quantitative data that can be evaluated statistical and be used for machine learning. Another option would be to do interviews with observers, which costs a lot more time and the data will have a qualitative character.

The experiment requires a set of paintings, large enough to train a machine learning algorithm. The size of the subject set is set to 15. The selected paintings must be abstract expressionistic paintings and have to be divers in their appearance. The size of the on screen images of the paintings must be high enough to reveal detail and small enough to be presented online in a browser window. The size of the screen images are set to be 600 pixels in height. The paintings are found at the Deviant Art web site, the largest artist's community in the world. These paintings are free to use with respect to their license.

The evaluation environment is a website within a browser window on a computer screen. The website contains no colors except shades of gray, this to minimize the diversion from the subject. The observer is guided through the website in the English or Dutch language and introduced to the questionnaire to be taken on each subject of the paintings.

The questionnaire uses the arousal and valence dimensions of the appraisal theory [ROSEMAN]. Examples on implementations on the arousal and valence dimensions are found at the tools Feeltrace [COWIE] and Geneva Emotion Wheel [SCHERER]. The dominance dimension is an addition to these dimensions, it is said to make it possible to distinct fear from anger [KEHREIN]. Therefore the questionnaire does also include the dominance dimension.

The design of the questionnaire on the emotion dimensions is as the following. The three emotion dimensions arousal, valence and dominance contain a rating on the evoked emotion. These dimensions can be rated on a 5 level scale from 1 to 5. The following table gives an overview of these dimensions, their scales and meaning.

Dimension	**Scale**	**Meaning**
Arousal	1 to 5	From passive to active From weak to strong
Valence	1 to 5	From unpleasant to pleasant From unattractive to attractive
Dominance	1 to 5	From submissive to dominant

Table 1 - emotion dimension, rating scales and their meaning

Besides rating the evoked emotion, the observer has an option to fill in a textual field to phrase the evoked feeling in words. This option is given in case the emotion dimensions are not sufficient to evaluate the evoked emotion. Also it gives the possibility to see if the emotion dimensions are understood by the human raters.

4. FEATURE EXTRACTION

Features are extracted from the images of the paintings. These features contain information about, for example colors and textures found in the image. Together the extracted features are called the feature vector. The used features are discussed below.

Before extracting features from the image, the image is indexed as a RGB image and a HSV image. An RGB indexed image contains information on the amount of red, green and blue of each pixel of the image. An HSV indexed image contains information on the hue, saturation and value of each pixel of the image.

Color

There is a range of color models to notate a color, e.g. RGB, HSV, HSL, CMYK and CIE XYZ. The RGB and CMYK models are used for technical reproduction of colors. The CIE XYZ model is based on direct measurements of the human eye, and serves as the basis for the HSV and HSL models. The HSV and HSL models closely model the human experience of a color and are useful tools for artists.

According to the HSV model color holds information on hue, saturation (also known as chroma or purity) and value. Hue is the part of the color described with names as yellow, red, blue, etc. Saturation is the intensity of a specific hue. Colors with a high saturation value are seen as vivid colors and colors with a low saturation value are seen as a shade of grey. The value tells if a color is dark, light or somewhere in between.

Color histograms are used as a feature on the color distribution of an image. Histograms derived from a RGB indexed image contain histograms on the red, green and blue distributions. A HSV indexed image contain histograms on the hue, saturation and value distributions. These histograms contain 10 bins, this number is gained by comparing a 256 bins histogram with histograms of a decreased number of bins. The 10 bins histogram still shows the same peaks as a 256 bins histogram.

Statistics and histograms on both RGB and HSV models are included in the feature vector used. This because the HSV model closely models the human experience of colors and the RGB model is related to the three color receptors in the human eye.

Texture

The three approaches used to describe texture are statistical, structural and spectral. Of these approaches only the statistical one is used, because it is available within the Matlab application.

Statistical techniques calculate texture by the statistical features of the grey levels of a image. These properties are computed from the grey level histogram or grey level co-occurrence matrix of the surface. Features often used are entropy, correlation, energy, contrast (inertia) and entropy homogeneity [HARALICK], these are explained below.

- Entropy, a measurement on a color distribution that gives a value on the amount of information the distribution contains. Images with a constant value on a color feature have an entropy of zero bits on the respective feature. Images that have a random distribution on a color feature have a high entropy on the respective feature.
- Features that are derived from the grey level coocurrence matrix (GLCM). The GLCM is calculated from an indexed intensity image on a diagonal of 0, 45, 90 and 135 degrees. Because of the symmetric method used the diagonals of 180, 225, 270 and 315 degrees give the same GLCM as the 0, 45, 90 and 135 degrees GLCM respectively. These features are:

- Correlation, a measure of how correlated a pixel is to its neighbor over the whole image. Correlation is 1 or -1 for a perfectly positively or negatively correlated image.
- Energy, gives the sum of squared elements in the GLCM. Energy is 1 for a constant image.
- Contrast, a measure of the intensity contrast between a pixel and its neighbor over the whole image. Contrast is 0 for a constant image.
- Entropy homogeneity, a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Homogeneity is 1 for a diagonal GLCM.

Structural techniques calculate texture as being composed of simple texture elements, that are regularly arranged on a surface according to a set of grammar rules.

Spectral techniques are based on properties of the Fourier spectrum and describe global periodicity of the grey levels of a surface by identifying high energy peaks in the spectrum.

Shape

A shape is obtained by characterizing the border of the shape by for example Fourier descriptors, invariant moments, shape measures, skeletons and edge abruptness.

Shape does not refer to the shape of an image but to the shape of a particular region within the image. Shapes will often be determined by first applying segmentation or edge detection to an image. In some cases accurate shape detection will require human intervention because methods like segmentation are very difficult to completely automate.

Algorithms that can be used for edge detection are Canny [CANNY] and Sobel. These algorithms provide a grayscale or monochrome image with the edge found in the original image. Statistics as mean and standard deviation on the grayscale images are included in the feature vector.

Feature vector

Because at this point it is uncertain which features will perform better than others, al the features described above are included to the feature vector for now. The feature vector contains the following features:

- Statistical features
  - A 10 bins histogram on the red distribution of RGB indexed pixels
  - A 10 bins Histogram on the green distribution of RGB indexed pixels
  - A 10 bins Histogram on the blue distribution of RGB indexed pixels
  - A 10 bins Histogram on the hue distribution of HVS indexed pixels
  - A 10 bins Histogram on the saturation distribution of HVS indexed pixels
  - A 10 bins Histogram on the value distribution of HVS indexed pixels
  - Entropy of the red distribution of RGB indexed pixels
  - Entropy of the green distribution of RGB indexed pixels
  - Entropy of the blue distribution of RGB indexed pixels
  - Entropy of the hue distribution of HSV indexed pixels

    - Entropy of the saturation distribution of HSV indexed pixels
    - Entropy of the value distribution of HSV indexed pixels
    - Mean of the red distribution of RGB indexed pixels
    - Mean of the green distribution of RGB indexed pixels
    - Mean of the blue distribution of RGB indexed pixels
    - Mean of the hue distribution of HSV indexed pixels
    - Mean of the saturation distribution of HSV indexed pixels
    - Mean of the value distribution of HSV indexed pixels
    - Standard deviation of the red distribution of RGB indexed pixels
    - Standard deviation of the green distribution of RGB indexed pixels
    - Standard deviation of the blue distribution of RGB indexed pixels
    - Standard deviation of the hue distribution of HSV indexed pixels
    - Standard deviation of the saturation distribution of HSV indexed pixels
    - Standard deviation of the value distribution of HSV indexed pixels
- Texture features on the hue, saturation and value part of HSV indexed pixels
    - Crontrast derived from the GLCM at 0, 45, 90 and 135 degrees
    - Correlation derived from the GLCM at 0, 45, 90 and 135 degrees
    - Energy derived from the GLCM at 0, 45, 90 and 135 degrees
    - Homogeneity derived from the GLCM at 0, 45, 90 and 135 degrees
- Edge features on the hue, saturation and value part of the HSV color distribution
    - Mean of the grayscale distribution of the Sobel edge enhanced image on the horizontal, vertical and both directions
    - Mean of the grayscale distribution of the Canny edge enhanced image

5. DATA EVALUATION AND CLEANING

The gathered data on the evoked emotions is evaluated on the following statistic measurements:

- Agreement on the paintings between observers on each of the emotion dimensions
- Histograms of rating distributions on each of the emotion dimensions
- Correlation between evaluated emotion dimensions
- Comparison of the subjective adjective given by raters in relation with the emotion dimensions
- Correlation between images features and the emotion dimensions

Agreement is calculated by using the Fleiss' kappa statistics [FLEISS, COHEN]. A measured agreement above 0 gives an indication that the agreement between observers is above chance level. A measured agreement of 1 indicates a perfect agreement between observers.

The correlation between the evaluated emotion dimensions tells something about the distinction between these dimensions as they are evaluated by human raters.

The histograms of rating distributions of the emotion dimensions will give information on the paintings if they are rated as evoking distinct emotions or not.

The subjective adjectives given in the textual field are compared to the data from the rated emotion dimensions. This will tell something about the observers' interpretation of the meaning of the emotion dimensions. It can be that this meaning as interpreted by the observers differs from the actual meaning of these dimensions.

To find any relation between the features of abstract expressionistic paintings and the evoked feeling as evaluated, correlations between extracted features from the paintings and the evaluation data are examined. In case no correlation is found between the extracted features and the evaluation data, predicting the evoked emotion based on these features will be difficult. This can mean that the used features are wrong. Correlations that are found can give significant information about the effect the correlating features have on the evoked emotion. The correlating features will be used as input data for the machine learning algorithms.

This set of paintings is the same as is used for data acquisition. The corpus on evoked emotion is also the same gained by the data acquisition.

6. MACHINE LEARNING

A set of machine learning algorithms (classifiers) are used to predict the emotional value of a painting. A classifier is trained by the emotion corpus as gathered from the previous steps as target data and the feature vectors extracted from the paintings as input data. Evaluation is done on each of the arousal, valence and dominance emotion dimensions.

The outcome of the evaluation will tell something about the ability of a classifier to evaluate an abstract painting as humans do. This with respect to the agreement on the evaluation of evoked emotions between human observers. By comparing the results between human observers and the classifiers it is possible to conclude on the research question.

There are many machine learning methods available that apply to classification problems like this one. It is not certain which method will perform best as they depend largely on the characteristics of the data used. Van der Walt and Barnard did show this by investigating very specific artificial data sets to determine conditions under which certain classifiers perform better and worse than others [WALT]. Selecting a machine learning methods applicable to this problem is a matter of trial and error.

The program Weka [WITTEN] is used for evaluation of the classifiers. Some well known machine learning algorithms are applied; Bayes Net, Naïve Bayes, Multi Layer Perceptron, Random Forest, Naïve Bayes Tree, Decision Stump and Decision Stump with Ada Boost.

Training and evaluating the algorithms are done by n-fold cross-validation experiments, also known as leave-one-out cross-validation. This approach is often used when the size of the training set is relatively small. The classifier can be trained using the complete data set. The method evaluates the classifier n times, where n is the number of entries in the training set. The resulting estimations are the averages of n evaluations.

The majority voting of all human observers serves as hard reference to determine the accuracy (the percentage of correct classifications) of the classifiers. A correct classification on a emotion dimension for a painting holds the same predicted value as is gained by taking the mode of all ratings given by the observers on the same emotion dimension and painting.

Next to the accuracy of the classifiers, the MAE (mean absolute error) is measured. This gives an idea on how wrong a classifier is when it is wrong. The closer the MAE is to zero, the better the performance of the classifier. Usually this measurement only applies to regression problems, but because the classifiers have to predict a nominal value it is also applicable to this classification problem.

To increase the performance of the classifiers the size of the feature vector can be decreased in iterations. A large set of features in relation with the number of training data entries can be a cause of poor performance of the classifiers. The methods used to evaluate the features are Information Gain, Relief Attribute Evaluation, SVM Attribute Evaluation and Symmetrical Uncertain Attribute Evaluation from the WEKA application. The results from these methods give a ranking on each of the features. A number of best ranking features on each of the emotion dimensions are used in the next iteration of machine learning. This approach will result in a used feature set with only features that are relevant to an emotion dimension and thus an optimal performance.

A try-out is done on a set of images that are outside of the scope of abstract expressionistic paintings, for example a set of websites. A emotion corpus is build on this

set of images and this corpus is used as a test set on the best performing classifiers on each of the emotion dimensions. These classifiers are trained by the emotion corpus and image features of the set of abstract expressionistic paintings. The results are compared to the results found on the classification of the paintings. This is done to find out if the results found are also applicable outside the domain of abstract expressionistic paintings. It gives information about the interoperability of the classifiers trained with the emotion corpus from the abstract expressionistic paintings.

7. EXPERIMENT RESULTS

This chapter describes the results found by conducting the experiments and evaluating the gathered data. The topics discussed in this section are; 'Diversity of paintings', 'Evaluation of the emotion corpus', 'Correlation analysis' and 'Comparison of machine learning algorithms'.

7.1. DIVERSITY OF PAINTINGS

The input data is expected to be divers. The diversity can be found by analyzing the distributions of the extracted features. Features that are expected to hold information on the diversity of the paintings are; the mean of the HSV color attributes and the entropy of the color's value attribute. These distributions can be found in the appendicle "*Image features*".

The distributions on the features of input images show a divers distribution on the features. According to this data the set of input data can be seen as divers and be used for the next steps of the experiment.

7.2. EVALUATION OF THE EMOTION CORPUS

The results of the questionnaire make the emotion corpus, evaluation of this corpus is done in the following section.

Overall agreement

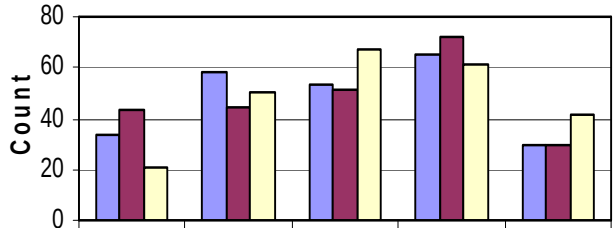
Measurement on agreement is done by applying the Fleiss Kappa calculation on the data. Agreements are calculated for the three evaluation scales; the valence, arousal and dominance dimensions. These dimensions all have 5 categories. Also displayed is the mean agreement as calculated by taking the mean of the percentage agreement values for each emotion dimension. See Table 3, Table 4 and Table 5 for the values of the percentage agreement on the paintings.

Scale	Fleiss Kappa	Mean agreement
Arousal	0.098	0.294
Valence	0.014	0.245
Dominance	0.070	0.295

Table 2 - kappa values on 5 categories

The results show a slight agreement between the raters on the dimensions arousal and dominance. The valence scale shows a poor agreement between raters. There is a notable difference between the Fleiss Kappa values and the mean agreement. This can be explained, the Fleiss Kappa measures the extent of the amount of agreement among observers to what would be expected if all observers made their ratings completely randomly. So when using 5 categories the agreement expected if all raters made their ratings completely random is 0.2. For the arousal category the extend given by Fleiss Kappa is 0.098 which makes sense given the mean agreement is 0.294.

The distribution of the ratings on these scales is shown below. For the valence dimension the ratings 1 and 5 are less often used by the raters. These ratings can possible be seen as extreme ratings and do not often apply to the images or not preferred by the raters. The same is also true for the arousal and dominance dimension. Therefore one can argue that a rating of 2 equals a rating of 1 and a rating of 4 equals a rating of 5.



Rating	1	2	3	4	5
Valence	34	58	53	65	30
Arousal	43	44	51	72	30
Dominance	21	50	67	61	41

Figure 1 - distribution of ratings

Agreement per painting

The following tables display the agreement on each painting on each emotion dimension. Displayed is the percentage agreement, the mean and the standard deviation as the result from the emotion corpus. For each of the emotion dimensions it shows a moderate agreement on each of the paintings.

Arousal

Painting	Agreement	Mean	StdDev
Post Consumer by Erkonom.	0.21	2.67	1.25
Rainforest by Blakewood	0.22	2.95	1.18
Open Mind by nighty	0.22	2.51	1.17
Europa by greycom	0.25	2.23	1.19
Elctricwiresscrossroads by easyfloat	0.25	3.13	1.11
Orange Burn by Violetlove	0.26	2.56	1.08
Red Boat by Eyepegus	0.27	3.41	1.13
This Kiss Could Feed A Family by TomSoya	0.28	3.59	0.90
Untitled 16 by Arthurx7	0.29	2.03	0.89
Sunset by Artbyelm	0.32	1.92	1.07
OMINOUS by Loci	0.33	4.10	0.78
No 21	0.33	1.92	1.05
Colorblind by Urban Paradox	0.35	3.44	0.93
Other II by Jarmen	0.41	3.82	0.81
The Trees by Danmorelle	0.42	4.31	0.76

Table 3 - agreement and distribution on the arousal scale

Valence

Painting	Agreement	Mean	StdDev
Post Consumer by erkonom	0.21	3.13	1.22
Orange Burn by Violetlove	0.21	3.03	1.21
Sunset by Artbyelm	0.21	3.05	1.18
Europa by Greycom	0.21	2.74	1.41
OMINOUS by Loci	0.23	3.36	1.12
Colorblind by Urban Paradox	0.23	2.79	1.14
The Trees by Danmorelle	0.24	3.23	1.10
No 21	0.24	3.33	1.09
Other II by Jarmen	0.26	2.62	1.12
Open Mind by Nighty	0.26	3.21	1.20
Elctricwiresscrossroads by Easyfloat	0.26	3.41	1.08
Untitled 16 by Arthurx7	0.26	3.41	0.98
This Kiss Could Feed A Family by TomSoya	0.27	2.64	1.03
Rainforest by Blakewood	0.29	3.41	1.17
Red Boat by Eyepegus	0.30	2.95	0.85

Table 4 - agreement and distribution on the valence scale

Dominance

Painting	Agreement	Mean	StdDev
Sunset by Artbyelm	0.21	2.92	1.21
Open Mind by Nighty	0.21	2.49	1.30
Orange Burn by Violetlove	0.22	2.79	1.14
Europa by Greycom	0.25	2.90	1.22
No 21	0.25	2.41	1.10
Rainforest by Blakewood	0.27	3.31	0.94
Electricwiresscrossroads by Easyfloat	0.27	3.26	0.93
Untitled 16 by Arthurx7	0.28	2.31	0.96
Post Consumer by Erkonom	0.29	2.51	0.87
Red Boat by Eyepegus	0.31	3.72	0.81
Colorblind by Urban Paradox	0.32	3.44	0.87
This Kiss Could Feed A Family by TomSoya	0.33	3.92	0.89
OMINOUS by Loci	0.39	3.90	0.67
Other II by Jarmen	0.41	3.56	0.81
The Trees by Danmorelle	0.42	4.10	0.78

Table 5 - agreement and distribution on the dominance scale

When comparing each of the emotion dimensions, the arousal and dominance dimension both have a similar lowest and highest agreement. The lowest agreement on the valence dimension is just as low as for the other, but the highest agreement is lower. Agreement between raters does not only depend on the emotion dimension as suggested by the kappa results but also on the painting self.

Subjective descriptive words used by raters

The textual evaluation results hold qualitative information. These are compared to the ratings on the emotion dimensions as evaluated on the same painting. This will give an idea if the dimensions where misunderstood by the observers. The possibility to use words was optional, so used when the observer needed more than the evaluation scales to evaluate the felt emotion. All the words are in Dutch. An arousal value of 1 means very passive and 5 very active, a valence value of 1 means not attractive and 5 means very

attractive, a dominance value of 1 means submissive and 5 means dominant. The words are given in order of their frequency.

Arousal

1	rust, stilte, kalm, dromerig, elegant, melancholisch, agressief, druk
2	herfst, droomeffect, verschillige rust, warmte, energie, beetje organisch, samenhangend onverschillig, te veel door elkaar, vaag, subtiel, rustgevend, strak, evenwicht, simpelheid, obscure, controle
3	passief levendig, ruige achtergrond, trieste vrolijkheid, rustige warmte, zoekend, saai, agressief, natuur, rommeltje, depressief, destructief, evenwicht
4	druk, levendig, ongeordend, chaos, agressief, zooitje, kermis, frisheid, warrig, duizelig, warm, passievol
5	gejaagd, fantasie, vrij, belegering, stress, agressie

Valence

1	vind ik niks, niet echt wat, te vaag, te rommelig, onlogisch, deprimerend, niets zeggend, koud, somber, destructief, te druk, ongeordend, kermis, bah
2	agressief, verschillig, vleeswond, niet bijzonder mooi, strak, rommeltje, zooitje, vaag, warrig, nietszeggend
3	beetje melancholisch, lekker rustig, belemmerd, niet indrukwekkend, evenwicht, controle, trieste vrolijkheid, beetje saai, briljant eigenlijk
4	wel mooi, aantrekkelijk, beschadiging, leuk, geweldige uitdrukking, wel tof, vrolijk, aandacht, frisheid, warm, passievol, vrij
5	elegant, aangenaam, vet, mooi, warmte, prachtig, warmte, stressvol

Dominance

1	doet mij niet veel, verschillig
2	vaag, rustgevend, rommeltje, warme rust, lekker rustig, controle, beschadiging, lichte kalmte, lichte rust, te abstract, elegant, aangenaam, droomeffect, subtiel
3	niets zeggend, zoekend, agressief, warrig, evenwicht, sleur, herfst, passief levendig, natuur
4	onlogisch, chaos, agressief, druk, slecht voor humeur, bosbrand, briljant, belegering, duisternis, onverschillig, obscuur, kapot, hoge bomen, warm, passievol, stress, prachtig
5	duivel aan een kruis, gejaagdheid, agressie, vleeswond, melancholisch, aandacht, 3D

Table 6 - words related to emotion dimensions

Overall the emotion dimensions and the related words do not point at a misunderstanding of the emotion dimensions. Although for example the words 'agressief' (aggressive) and 'druk' (busy) are used in relation with a low arousal value, words as 'rust' (peace) and 'stilte' (quietness) are used more often in the same relation.

Correlation between emotion dimensions

The evaluation results are analyzed to find correlation between the emotion dimensions. This is done to find out if the observers interpreted the emotion dimensions as distinct

dimensions. The tables below show the correlation between the mode, mean and standard deviation of the results within the emotion corpus.

Mode	Arousal	Valence	Dominance
Mode arousal	1.00	0.67	**0.00**
Mode valence	0.67	1.00	0.57
Mode dominance	**0.00**	0.57	1.00
Mean arousal	**0.00**	0.89	**0.00**
Mean valence	0.73	**0.00**	0.18
Mean dominance	**0.00**	0.61	**0.00**
Stdev arousal	0.11	0.89	**0.03**
Stdev valence	0.08	0.32	0.15
Stdev dominance	**0.00**	0.50	**0.00**

Table 7 - P-values on correlation between the mode and statistics of target data

Mean	Arousal	Valence	Dominance
Mode arousal	**0.00**	0.73	**0.00**
Mode valence	0.89	**0.00**	0.61
Mode dominance	**0.00**	0.18	**0.00**
Mean arousal	1.00	0.46	**0.00**
Mean valence	0.46	1.00	0.27
Mean dominance	**0.00**	0.27	1.00
Stdev arousal	**0.01**	0.70	**0.03**
Stdev valence	0.27	0.62	0.24
Stdev dominance	**0.00**	0.97	**0.00**

Table 8 - P-values on correlation between the mean and statistics of target data

Standard deviation	Arousal	Valence	Dominance
Mode arousal	0.11	0.08	**0.00**
Mode valence	0.89	0.32	0.50
Mode dominance	**0.03**	0.15	**0.00**
Mean arousal	**0.01**	0.27	**0.00**
Mean valence	0.70	0.62	0.97
Mean dominance	**0.03**	0.24	**0.00**
Stdev arousal	1.00	0.20	**0.02**
Stdev valence	0.20	1.00	0.05
Stdev dominance	**0.02**	0.05	1.00

Table 9 - P-values on correlation between the standard deviation and statistics of target data

The results show a strong correlation between the arousal and dominance dimensions. Also there is a strong correlation between the standard deviation of the dominance dimension and the mode and mean of the arousal and dominance dimensions. Another strong correlation is found between the standard deviation of the arousal dimension and the mode and mean of the dominance dimension.

This means the arousal and dominance dimensions are interpreted in more or less the same way by the raters. The valence dimension is interpreted in a different way as the arousal and dominance dimensions.

The correlation between the standard deviation of the arousal and dominance dimensions and their mean and mode values shows that the agreement between observers is higher when the painting is evaluated as highly arousing or highly dominant.

7.3. CORRELATION ANALYSIS

Correlation analysis is done between the extracted features of a painting and evaluation data on the emotion dimensions. This is done to find relations between features of a painting and the emotion dimensions. All the extracted image features from the previous proposed feature vector are analyzed. The tables below show the results on the p-values (the strength) of their correlations.

Entropy of color distributions

Entropy	Red	Green	Blue	Hue	Sat.	Value
Mode arousal	0.13	0.12	**0.01**	0.20	0.32	**0.03**
Mode valence	0.21	0.21	0.44	0.57	0.55	0.34
Mode dominance	0.55	0.22	0.10	0.31	0.64	0.20
Mean arousal	0.10	0.08	**0.01**	0.08	0.21	**0.04**
Mean valence	0.40	0.44	0.64	0.52	0.94	0.53
Mean dominance	0.06	**0.03**	**0.01**	**0.03**	0.07	**0.02**
Stdev arousal	0.43	0.63	0.41	**0.05**	0.65	0.32
Stdev valence	0.51	0.57	0.34	0.35	0.27	0.46
Stdev dominance	0.19	0.19	**0.01**	0.14	0.85	**0.05**

Table 10 - P-values on the correlation between entropy values of colors and emotion dimensions

The entropy value of a color distribution gives a number on the amount of information the distribution contains. The entropy on the green distribution shows a relation with the mean of the dominance dimension. The entropy on the blue distribution shows a relation with the mode arousal, mean arousal, mean dominance and the standard deviation of the dominance dimension. The entropy on the hue distribution shows a relation with the mean of the dominance and the standard deviation of the arousal dimension. The entropy on the value distribution shows a relation with the mode and mean of the arousal, the mean and standard deviation of the dominance dimension.

Entropy on the color distributions seems to have a strong relation to the mean of the arousal and dominance dimensions. Hence arousal and dominance are related to the amount of information contained in the image.

Mean of color distributions

Mean	Red	Green	Blue	Hue	Sat.	Value
Mode arousal	0.81	0.82	0.98	0.73	0.61	0.45
Mode valence	0.62	0.27	0.42	0.61	0.57	0.26
Mode dominance	0.81	0.58	0.65	0.85	0.82	0.65
Mean arousal	0.26	0.50	0.94	0.19	0.71	0.19
Mean valence	0.43	0.17	0.35	0.99	0.92	0.23
Mean dominance	0.18	0.12	0.34	0.47	0.96	0.06
Stdev arousal	0.79	0.90	0.82	0.82	0.60	0.70
Stdev valence	0.38	0.48	0.58	0.37	0.94	0.45
Stdev dominance	0.76	0.50	0.38	0.58	0.15	0.82

Table 11 - P-values on the correlation between mean values of colors and emotion dimensions

The mean values of the colors in a painting show almost no correlation with the emotion dimensions. This means that the quantity of a color alone has almost no effect on the felt emotion.

Standard deviation of color distributions

Standard deviation	Red	Green	Blue	Hue	Sat.	Value
Mode arousal	0.06	**0.00**	**0.00**	0.14	**0.01**	**0.00**
Mode valence	0.88	0.91	0.75	0.09	0.89	0.95
Mode dominance	0.09	**0.00**	**0.02**	0.09	0.08	**0.01**
Mean arousal	0.07	**0.00**	**0.00**	0.18	**0.00**	**0.01**
Mean valence	0.99	0.99	0.56	0.11	0.57	0.91
Mean dominance	**0.03**	**0.00**	**0.00**	**0.02**	**0.00**	**0.00**
Stdev arousal	**0.03**	**0.03**	**0.04**	0.62	0.33	**0.02**
Stdev valence	0.49	0.29	0.36	0.47	0.15	0.41
Stdev dominance	0.13	**0.00**	**0.00**	0.29	0.16	**0.01**

Table 12 - P-values on the correlation between standard deviation of colors and emotion dimensions

The standard deviation on a color distribution tells something about the width of a distribution. The standard deviation on the hue distribution holds information about the range of colors used within an image. The standard deviation on the red, green and blue distributions do not tell us about the diversity of the colors in the image , but about the range of the intensity of these colors and thus about the range of the intensity within the image itself. So results on the correlation of the standard deviation of the red, green and blue distributions can also be seen on the value distribution.

The standard deviation on the red, green and blue distributions show a relation with the mode, mean and standard deviation of the arousal and dominance dimensions. This is also true for the standard deviation on the value distribution. The standard deviation on the saturation distribution relates with the mode and mean of the arousal and dominance dimensions, but not with the standard deviation of these dimensions. The standard deviation on the hue distribution does relate to the mean of the dominance dimension.

These correlations show that the width of a color distribution relates to the arousal and dominance dimensions.

Color histograms

The results of the correlation analysis on the RGB and HSV histograms can be found in the appendix of this document. Because a histogram contains a number of bins and for correlation analysis these bins are analyzed separately, absence of correlation does not mean a histogram in total has no relation to the emotion dimensions. Correlations on the arousal dimension are found in the red part of the RGB distribution and the hue part of the HSV distribution. Correlations on the valence dimension are found in the green part of the RGB distribution and the hue part of the HSV distribution. Correlations on the dominance dimension are found in the red and green parts of the RGB distribution and the value part of the HSV distribution.

Gray level co-occurrence matrices

The results of the correlation analysis on statistics derived from the gray level co-occurrence matrices (GLCM) can be found in the appendix of this document. Statistics derived from GLCM's are used to describe texture features of an image. The GLCM's are calculated with a given direction (0, 45, 90 and 135 degrees) on the parts of the HSV color distribution. From these GLCM's statistics are derived, contrast, correlation, energy and homogeneity. This results in a total of 48 features. Correlations are found on the arousal and dominance dimensions and only few on the valence dimension. The standard deviation on the valence dimension correlates with the correlation feature on the value part of the HSV color distribution. On the arousal and dominance dimension correlations are found with the contrast, energy and homogeneity features on the saturation and value parts of the HSV color distribution. The hue part of the HSV color distribution does only correlate with the dominance dimension, on the features contrast and energy.

Edges

Correlation analysis between the Sobel and Canny edge features on the parts of the HSV color distribution and emotion dimensions can be found in the appendix of this document, it gives the following results. The Sobel edge features on the hue distribution do not show any correlations with the emotion dimensions. Correlations are found between the mean of the arousal and dominance dimensions and the Canny edges of saturation and value distributions. Extracting the Sobel edge features in a horizontal or vertical direction does not add any correlations opposed the features extracted in both directions. Sobel edge features extracted in both directions do show correlation with the standard deviation of the dominance dimension. The Canny edge features do not show any correlation. The differences in the amount of correlating features is probably because the Canny algorithm tends to have a higher recall than the Sobel algorithm and produces more dense images. It is possible that other statistics on the images produced by the Canny algorithm do show other correlations.

Totals

When counting the number of correlating features with the mode, mean and standard deviation of the emotion dimensions, the following table is found. It shows the highest number of correlating features for the arousal and dominance dimensions.

	Num. of correlating features
Mode arousal	33
Mode valence	13
Mode dominance	4
Mean arousal	29
Mean valence	3
Mean dominance	50
Stdev arousal	11
Stdev valence	9
Stdev dominance	15

Table 13 - numbers of correlating features

The table shows a high number of features that correlate with the arousal and dominance dimensions and less with the valence dimension. It is possible there are less image features that relate to the valence emotion dimension in general or just that most of the extracted features are related to the arousal and dominance dimensions. The number of features related to the standard deviations of the emotion dimensions show that there is a possible relation of a number of image features and agreement between raters.

Correlating feature vector

The results of the correlation analysis between the extracted features of a painting and evaluation data on the emotion dimensions show the features that are somehow related to the emotion dimensions, only these features will be used for machine learning. The feature vector as it will be used for machine learning contains the following features:

- Statistical features
    - A 10 bins histogram on the red distribution of RGB indexed pixels
    - A 10 bins Histogram on the green distribution of RGB indexed pixels
    - A 10 bins Histogram on the blue distribution of RGB indexed pixels
    - A 10 bins Histogram on the hue distribution of HVS indexed pixels
    - A 10 bins Histogram on the saturation distribution of HVS indexed pixels
    - A 10 bins Histogram on the value distribution of HVS indexed pixels
    - Entropy of the red distribution of RGB indexed pixels
    - Entropy of the green distribution of RGB indexed pixels
    - Entropy of the blue distribution of RGB indexed pixels
    - Entropy of the hue distribution of HSV indexed pixels
    - Entropy of the saturation distribution of HSV indexed pixels
    - Entropy of the value distribution of HSV indexed pixels
    - Standard deviation of the red distribution of RGB indexed pixels
    - Standard deviation of the green distribution of RGB indexed pixels
    - Standard deviation of the blue distribution of RGB indexed pixels
    - Standard deviation of the hue distribution of HSV indexed pixels

    - Standard deviation of the saturation distribution of HSV indexed pixels
    - Standard deviation of the value distribution of HSV indexed pixels
- Texture features on the hue part of HSV indexed pixels
    - Contrast derived from the GLCM at 0, 45, 90 and 135 degrees
    - Energy derived from the GLCM at 0, 45, 90 and 135 degrees
- Texture features on the saturation part of HSV indexed pixels
    - Contrast derived from the GLCM at 0, 45, 90 and 135 degrees
    - Energy derived from the GLCM at 0, 45, 90 and 135 degrees
    - Homogeneity derived from the GLCM at 0, 45, 90 and 135 degrees
- Texture features on the value part of HSV indexed pixels
    - Crontrast derived from the GLCM at 0, 45, 90 and 135 degrees
    - Correlation derived from the GLCM at 0, 45, 90 and 135 degrees
    - Energy derived from the GLCM at 0, 45, 90 and 135 degrees
    - Homogeneity derived from the GLCM at 0, 45, 90 and 135 degrees
- Edge features on the saturation and value part of HSV indexed pixels
    - Mean of the grayscale distribution of the Sobel edge enhanced image on the horizontal, vertical and both directions

7.4. COMPARISON OF MACHINE LEARNING ALGORITHMS

The program Weka is used for machine learning and classification. Some well known machine learning algorithms are used; Bayes Net, Naïve Bayes, MultiLayer Perceptron, Random Forest, Naïve Bayes Tree and Decision Stump with AdaBoost.

The configuration parameters of these classifiers can be found in the appendix "Weka classifiers and configuration".

As test mode a 15 fold cross-validation is used, each classifier is tested 100 times, this results in 1500 classifications on each emotion dimension per classifier.

To increase the accuracy the size of used feature vector is decreased in iterations. The first vector contains 114 features. In relation with the size of the corpus, this can be a cause for poor accuracy of the classifiers. The methods used to evaluate the features are Information Gain, Relief Attribute Evaluation, SVM Attribute Evaluation and Symmetrical Uncertain Attribute Evaluation from the WEKA application. The results from these methods give a ranking on each of the features. The overall best ranking features on each of the emotion dimensions are used in the next iteration of machine learning. The results can be found in the appendix "Feature selection on the emotion dimensions".

The algorithms are trained with the correlating and evaluated features extracted from the set of abstract expressionistic paintings and the observer's rankings on emotion dimensions. This results in training the classifiers with 114, 20, 10, 5, 4, 3, 2 and 1 features on 15 painting. Performance is measured on accuracy and mean absolute error (MAE) of the classifiers on each emotion dimension. The results can be found in the appendix "Machine learning results".

Accuracy is measured by the percent of correct classifications. The MAE is computed by summing the absolute difference between the actual and predicted target value for each instance and then taking the average.

The results are discussed for each of the emotion dimensions, arousal, dominance and valence. The results are discussed for classifiers trained with the best performing number of features.

On the arousal emotion dimension

The result on the arousal dimension show that most classifiers have an accuracy around 50% to 70%. The classifiers with the best accuracy are; the Naive Bayes classifier trained with 2 features (accuracy 73.33%, MAE 0.16) and the Decision Stump with Ada Boost classifier trained with 1 feature (accuracy 73.33%, MAE 0.27). The Naïve Bayes classifier shows a decrease on performance when leaving out the feature '*standard deviation on the green part of the RGB distribution*' where the other classifiers show an increase. The overal best performing classifier is the Naive Bayes classifier trained with the features '*Sobel edge in both directions on the value part of the HSV distribution*' and '*standard deviation on the green part of the RGB distribution*'.

On the dominance emotion dimension

The results on the dominance dimension show a highest accuracy of 46.67% for the classifier Naive Bayes. The rest of the classifiers score below 40% on accuracy. The MAE shows the best performance for the Naive Bayes classifier. The Naïve Bayes classifier trained with 5 and 3 features both reach an accuracy of 46.67%, the MAE on 5 features is 0.28 and on 3 features is 0.29. This makes the Naive Bayes classifier trained with 5 features the best performing classifier on the dominance dimension (accuracy 46.67%, MAE 0.28). The used features are; '*standard deviation on the saturation part of the HSV distribution*', '*Sobel edge in both directions on the value part of the HSV distribution*', '*standard deviation on the green part of the RGB distribution*', '*energy of the GLCM on 135 degrees on the hue part of the HSV distribution*' and '*standard deviation on the blue part of the RGB distribution*'.

On the valance emotion dimension

The results on the valence emotion dimension show a highest accuracy of 73.33% for the classifiers Naïve Bayes Tree trained with 1 feature and Decision Stump with Ada Boost trained with 10 to 1 features, the other classifiers all have an accuracy below 60%. On the MAE measures the best performing classifiers are Naive Bayes trained with 4 and 3 features (MAE 0.25) and Random Forest trained with 3 features (MAE 0.25), followed by Naive Bayes Tree trained with 1 feature (MAE 0.28). This makes the Naïve Bayes Tree trained with the feature '*contrast of the GLCM on 90 degrees on the value part of the HSV distribution*' the preferred classifier for the valence emotion dimension.

On computation costs

Computation costs or efficiency of the classifiers is measured by the number of instructions it uses (CPU time) to train and test the classifier. The following table shows the results on 114 features.

Classifier	CPU time
Bayes Net	0.04001706
Naive Bayes	**0.00432605**
Multilayer Perceptron	1.10086279
Random Forest	0.57783088
Naïve Bayes Tree	0.07789188
Decision Stump with Ada Boost	0.03780720

Table 14 – classifiers and computation costs

Compared to human observers

To compare the classifiers algorithms to human raters, the overall accuracy and mean absolute error of the human raters is calculated on the corpus. They are found in the table below.

Emotion dimension	Accuracy	Mean absolute error
Arousal	41.71%	0.86
Dominance	41.71%	0.78
Valence	37.44%	0.97

Table 15 - accuracy and MAE on human raters

Compared to the classifier algorithms the human raters have a lower accuracy on all dimensions, also the MAE is much higher than that of the classifier algorithms.

The Naive Bayes classifier show the best result on the arousal and dominance dimensions, also this classifier is the most efficient one. The best performing classifier on the valence dimesion is the Naïve Bayes Tree, it is the forth most efficient classifier.

8. AN EXAMPLE ON THE EMOTION CLASSIFICATION OF WEBSITES

This experiment relates to the methodology described in the section Machine learning, it gives information about the applicability of the used approach outside of the domain abstract expressionistic paintings.

The results from the comparison of machine learning algorithms are tried on a set of websites. A total of 9 websites are rated by 2 human observers, the results from these observations are used as a test set for the classifiers. The classifiers are trained by the emotion corpus and image features gained from the abstract expressionistic paintings. This gives a training set containing 15 instances and a test set containing 9 instances on each of the emotion dimensions. The results of these test can be found below.

Emotion dimension	Accuracy	Mean absolute error
Arousal	33.33%	0.2656
Dominance	22.22%	0.3210
Valence	0.00%	0.3466

Table 16 - accuracy and MAE of emotion classification of websites

The results show that the classifiers do not perform as well on websites as on the paintings. The accuracy is only just above chance level on the arousal dimension and very low compared to the results found for the paintings. Also the MAE results are higher for the websites than for the paintings on each emotion dimension. Judging on these results the trained classifiers are not applicable to classifying the emotion dimensions of websites.

To gain more insight on the performances of these classifiers on the websites, the confusion matrices of these classifiers are shown below.

Emotion dimension															
Arousal					Dominance					Valence					
a	b	c	d	e	a	b	c	d	e	a	b	c	d	e	classified as
0	0	0	0	0	0	0	2	0	0	0	0	1	1	0	a = 1
3	0	0	0	0	0	0	2	0	0	0	0	0	3	0	b = 2
1	0	0	1	0	0	0	2	1	0	0	0	0	3	0	c = 3
0	0	0	3	0	0	0	2	0	0	0	0	0	0	0	d = 4
0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	e = 5

Table 17 - confusion matrices of emotion classification of websites

The confusion matrix of the arousal emotion dimension shows that the classifier often classified a class *b or c* as a class *a* and a class *c, d or e* as a class *d*. This shows the ability of the classifier to predict if an image of a website should be classified as a low arousing image or a high arousing image. The classifiers' confusions matrices of the dominance and valence dimensions show a tendency to assign each class to one class. Leaving the classifications on the dominance and valence dimensions meaningless. It is possible the image features used are not applicable to this domain and other image features are more meaningful to predict the evoked emotion. It is also possible that better results are gained when an emotion corpus gained by evaluating the websites is used to train the classifiers. To conclude on this experiment, knowledge gained by evaluating expressionistic paintings is not applicable outside its domain using a machine learning approach, although to some extend it is applicable to the arousal emotion dimension.

9. CONCLUSIONS AND RECOMMEDATIONS

This research on a machine learning approach to emotion classification of abstract expressionistic paintings involved human observers, feature extraction on images and classification algorithms. A number of research question are stated earlier, conclusions and recommendations are discussed here.

The test results on human observations show a minor agreement between human raters. The emotion evoked at these observers tend to differ a lot between observers. This makes it harder for a classifier to be precise, in this experiment classifiers are trained by a majority vote of all human observations. However none of the human observations are wrong, they are valid for the observer. Therefore classifiers should be trained taking the whole distribution of observations in account. Also, the machine learning algorithms could be used to predict a distribution instead of a nominal value. Predicted a distribution on an emotion dimension is more confirm the reality.

The different emotion dimensions evaluated are; arousal, dominance and valence. Arousal and dominance show a strong correlation to each other within the observations. Valence is not correlated with these other dimensions. Therefore arousal and dominance are related dimensions and valence is a distinct dimension. Also there is a distinction at the difference in agreement between observers, they agree more on the arousal and dominance dimensions than on the valence dimension. Another item is the correlation between the standard deviation of the arousal and dominance dimension, people tend to agree more on highly dominant and highly arousing paintings.

Correlation analysis on image features extracted from the paintings show many correlating features for the arousal and dominance dimensions and less for the valence dimension. Features extracted from the hue part of an HSV indexed image do less often correlate with the emotion dimensions than features extracted from the saturation and value parts, color by itself does not tend to relate to the emotion dimensions. Only the dominance dimension does correlate with some texture features derived from hue. In relation to the emotion dimensions saturation and value seem to be the most important aspects of an image.

The correlation analysis does also shows a number of correlating image features with the standard deviations of the emotion dimensions. These features can be used to predict the distribution of an emotion dimension on a painting.

Prediction of the evoked emotion by classification algorithms give a correct classification of 73.33% (MAE 0.16) on the arousal dimension, 46.67% (MAE 0.28) on the dominance dimension and 73.33% (MAE 0.28) on the valence dimension. Making the arousal dimension the easiest to predict an evoked emotion on. These results are found after applying feature selecting to reduce the number of features used to train the classifiers. This gives a number of features that are most valuable to predicting the emotion of a painting.

It shows that for the arousal emotion dimension the edges found within the value distribution and the diversity in intensity of the color green within the painting are most valuable. The more edges found in an image and the wider the distribution of the intensity of the color green the higher the arousal value.

For the dominance dimension most valuable are; the diversity in color saturation, the edges found within the value distribution, the diversity in intensity of the color green, texture energy derived from a 135 degrees GLCM within the hue distribution and the diversity in intensity of the color blue. A wider distribution of color saturation, more edges

in an image, a wider distribution of the intensity of the color green, lower texture energy and a wider distribution of the intensity of the color blue give a higher dominance value.

On the valence dimension the GLCM's contrast on 90 degrees within the value distribution, these are in some extend very small to very large horizontal lines in the image that cause to predict a higher value on the valence emotion dimension.

Compared to human raters the classifiers do perform better on the domain of abstract expressionistic paintings. For the domain of websites the classifiers performed significantly worse, performing just above chance level. This makes the results of this research not applicable outside the scope of abstract expressionistic paintings. The best results are gained on the arousal emotion dimension, it is in some extend possible to predict if an image of a website is perceived arousing or not.

To conclude on this research; of the three emotion dimensions examined, emotion on the arousal dimension is the easiest to predict. Edges found in the paintings' images are believed to be a fair predictor. Also diversity within the intensity of a color does seem to have a strong relation to emotion and color by itself does not. This in contradiction as what is commonly believed, e.g. to see red as an angry color and yellow as a happy color.

The dominance emotion dimension does not show very good results when predicting with the classifiers. Also most of its related image features are also used in relation to the arousal emotion dimension. It is questionable if the dominance dimension is a good measure for emotion in abstract expressionistic paintings.

The predicting feature on the valence emotion dimension needs more ground to be able to conclude on its relation.

The topic of emotion classification of abstract expressionistic paintings would need a better approach than the one described in this work. Distributions of human ratings should be used to train the classifiers and instead of predicting a nominal value, also a sort of distribution should be predicted. To achieve this a much larger emotion corpus is needed. All this because there is no true or false when asking a human observer to rate a painting on a number of emotion dimensions.

REFERENCES

[BATLINER] A. Batliner, S. Steidl, C. Hacker, E. Noth, and H. Niemann. 2005. Private Emotions vs. Social Interaction - towards New Dimensions in Research on Emotion. In Proceedings of a Workshop on Adapting the Interaction Style to Affective Factors, 10th International Conference on user Modelling, Edinburgh.

[BARTLETT] M. S. Bartlett, G. Littlewort, C. Lainscsek, I. Fasel, and J. Movellan. 2004. Machine learning methods for fully automatic recognition of facial expressions and facial actions. In Systems, Man and Cybernetics, 2004 IEEE International Conference on , vol.1, no., pp. 592-597 vol.1, 10-13 Oct.

[BURKHARDT] F. Burkhardt, J. Ajmera, R. Englert, J. Stegmann, and W. Burleson. 2006. Detecting Anger in Automated Voice Portal Dialogs. Interspeech. Berlin

[CANNY] J. Canny. 1986. A computational approach to edge detection. In IEEE Trans. on Pattern Analysis and Machine Intelligence. PAMI-8, no. 6, pp. 679-698

[CHOI] G. S. Choi, R. Oehlmann, H. Dlake, and D. Cottington. 2007. Cross-Cultural Comparison of Color Preferences between English and Korean Subjects. In Proceedings of the 6th workshop on social intelligence design.

[COHEN] J. Cohen. 1960. A coefficient of agreement for nominal scales. In Educational and Psychological Measurement 20: 37–46.

[COWIE] R. Cowie, E. Douglas-Cowie, S. Savvidou, E. McMahon, M. Sawey, and M. der Schro. 2000. FEELTRACE: An Instrument for Recording Perceived Emotion in Real Time. In Proceedings of the ISCA Workshop on Speech and Emotion, pp. 19–24. Belfast: Textflow.

[EKMAN] P. Ekman. 1988 An argument for basic emotions. In: Stein, N. L. and Collins,A. The Cognitive Structure of Emotions. Cambridge University Press.

[FLEISS] J. L. Fleiss. 1971. Measuring nominal scale agreement among many raters. Psychological Bulletin 76 (5): 378–382.

[HARALICK] R. M. Haralick, K. Shanmugam, and I. Dinstein. 1973. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybertinetics, SMC-3(6):610-621.

[HERIK] H. van den Herik. and E. Postma. Discovering the Visual Signature of Painters. Future Directions for Intelligent Systems and Information Sciences. The Future Speech and Image Technologies, Brain Computers, WWW, and Bioinformatics. Editor N. Kasabov. Physica Verlag (Springer-Verlag), pp. 129-147.

[HUANG] J. Huang, S.R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih. 1999. Spatial Color Indexing and Applications. Int'l J. Computer Vision, vol. 35, no. 3, pp. 245-268.

[ICOGLU] O. Icoglu, B. Gunsel, and S. Sariel. Classification and Indexing of Paintings Based on Art Movements, 12th European Signal Processing Conference 2004, September 6-10, 2004. Vienna, Austria, pp. 749-752.

[ITTEN] J. Itten. 1961. Kunst der Farbe (Studienausgabe).

[JACOBSON] N. Jacobson and W. Bender. 1996. Color as determined communication. IBM Syst. J., vol. 35, no. 3&4, pp. 526-538.

[KEHREIN] R. Kehrein. 2002. The prosody of authentic emotions. In: Proc. Speech Prosody Conference, pp. 423–426.

[KEREN] D. Keren. 2002. Painter Identification Using Local Features and Naïve Bayes. Proc. of the 16th International Conference on Pattern Recognition.

[KIRCH] J. Kirsch and R. Kirsch. The Anatomy of Painting Style: Description with Computer Rules. Leonardo, Vol. 21, No. 4, 1988, pp. 437-444.

[KRONER] S. Kröner and A. Lattner. 1988. Authentication of Free Hand Drawings by Pattern Recognition Methods. In Proc. of the 14th International Conference on Pattern Recognition.

[LYU] S. Lyu, D. Rockmore, H. Farid. 2004. A Digital Technique for Art Authentication. In Proceedings of the National Academy of Sciences of the United States of America, Vol. 101, No. 49, December 7.

[MATLAB] Matlab, http://www.mathworks.com/products/matlab/, accessed 08-05-2007

[MITCHELL] T. Mitchell. 1997. Machine Learning. WCB/McGraw-Hill.

[OCVIRK] O. G. Ocvirk, R. E. Stinson, P. R. Wigg, R. O. Bone, and D. L. Cayton. Art Fundamentals Theory & Practice Seventh Edition, Calmann & King LTD, 1994. ISBN 0-697-12545-9

[PETRUSHIN] V.A. Petrushin, Emotion recognition agents in real world, in AAA1 Full Symposium on Socially Inrelligent Agents: Human in the Loop, (North Falmouth, Massachusetts, 2000), pp. 136-138

[PLUTCHIK] Plutchik, R. The Psychology and Biology of `Emotion. Harper Collinns, New York, 1994

[RAO] A. Rao, R. K. Srihari, and Z. Zhang. 2000. Geometric histogram: A distribution of geometric configurations of color subsets. To appear in Proceedings of SPIE: Internet Imaging, 3964, January 2000.

[ROSEMAN] Roseman, Ira J. 1984. Cognitive Determinants of Emotion: A Structural Theory. In P. Shaver (Ed.) Review of Personality & Social Psychology, Vol. 5: Emotions, Relationships, and Health. (pp. 11 - 36). Beverly Hills, CA: Sage.

[RUSSEL] J. A. Russell. 1980. A Circumplex Model of Affect. In Journal of Personality and Social Psychology 39, pp. 1161-1178. Elsevier Science Inc.

[SABLATNIG] R. Sablatnig, P. Kammerer, E. Zolda. Hierarchical Classification of Paintings Using Face- and Brush Stroke Models. Vienna University of Technology, Institute of Automation, Pattern Recognition and Image Processing Group

[SCHERER] K. R. Scherer. 2005. What are emotions? And how can they be measured?", In Social Science Information, vol 44, no.4, pp. 695–729.

[SCHLOSBERG] H. Schlosberg. 1954. A scale for judgement of facial expressions. In Journal of Experimental Psychology 29,497-510.

[SMEULDERS] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. 2000. Content-Based Image Retrieval at the End of the Early Years, In IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec.

[SOBEL] I. Sobel, G. Feldman. 1968. A 3x3 Isotropic Gradient Operator for Image Processing. presented at a talk at the Stanford Artificial Project in 1968, unpublished but often cited, orig. in Pattern Classification and Scene Analysis, Duda,R. and Hart,P., John Wiley and Sons,'73, pp271-2

[STEIDL] S. Steidl, M. Levit, A. Batliner, E. Noth, H. Niemann. 2005. Of All Things the Measure is Man: Automatic Classification of Emotions and Inter-Labeler Consistency.In Proc.ICASSP 2005, Philadelphia.

[WALT] C. M. van der Walt, E. Barnard. 2006. Data characteristics that determine classifier performance. In Proceedings of the Sixteenth Annual Symposium of the Pattern Recognition Association of South Africa, pp.160-165.

[WANG04] W. Wang, Y. Yu, J. Zhang. 2004. Image emotional classification: static vs. dynamic. Systems, Man and Cybernetics. Volume 7: 6407- 6411

[WANG05] W. Wang, Y. Yu, J. Zhang. 2005. A new SVM based emotional classification of image. Journal of Electronics (China). Volume 22(1): 98-104.

[WIDJAJA] I. Widjaja, W. Leow, F. Wu. 2003. Identifying Painters from Color Profiles of Skin Patches in Painting Images. In Proc. of the IEEE International Conference on Image Processing, Barcelona, Spain.

[WITTEN] I. H. Witten, E. Frank. 2005. Data Mining: Practical machine learning tools and techniques, 2nd Edition, Morgan Kaufmann, San Francisco. [Online] Available: http://www.cs.waikato.ac.nz/~ml/weka/

APPENDICES

A. **Evaluation data**
The set of abstract expressionistic paintings used for this research

B. **Image features**
Provides an overview of a set of features extracted from the paintings' images

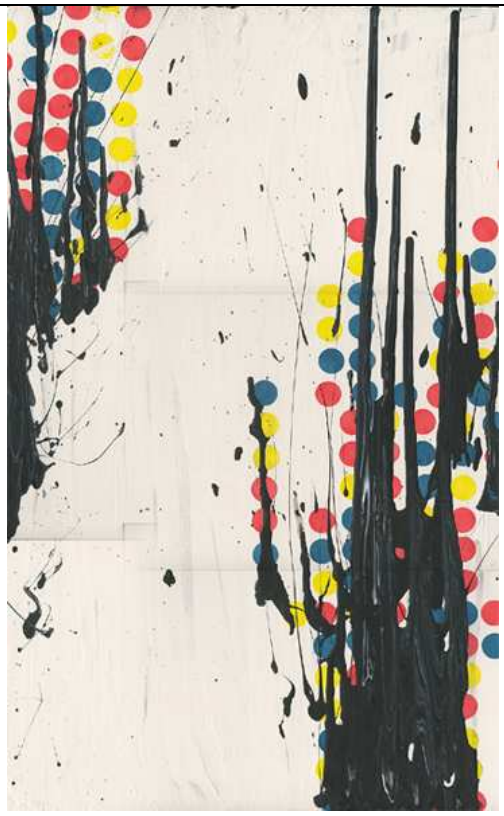
C. **Correlation analysis**
The results from the correlation analysis done on the paintings' features

D. **Feature selection**
A list of the paintings' features ranked on relevance to the emotion dimensions

E. **Classifiers and configuration**
The configurations used for the classifiers in this research

F. **Machine learning results**
The results of a set of classifiers on their ability to classify the emotion dimensions of the paintings

Evaluation data



Colorblind by Urban Paradox

Rainforest by Blakewood

Elctricwiresscrossroads by Easyfloat

Europa by Greycom

No 21

OMINOUS by Loci

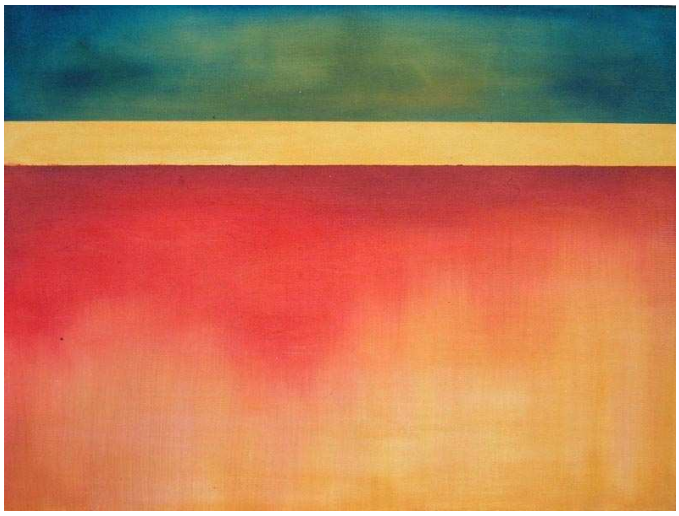
Open Mind by Nighty

Orange Burn by Violetlove

Other II by Jarmen

Red Boat by Eyepegus

Post Consumer by Erkonom



Sunset by Artbyelm

The Trees by Danmorelle

This Kiss Could Feed A Family by TomSoya

Untitled 16 by Arthurx7

Image features

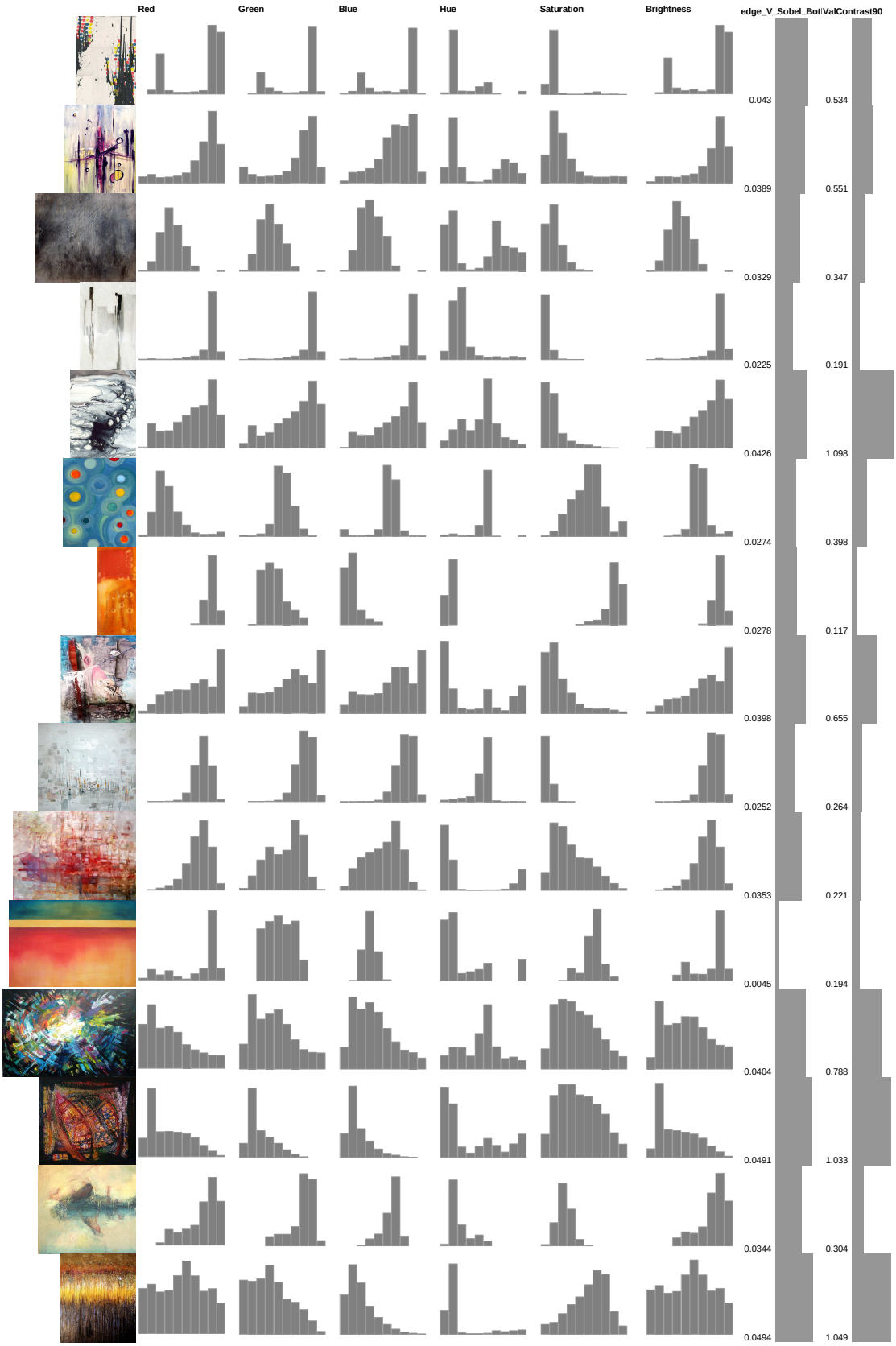


Image features

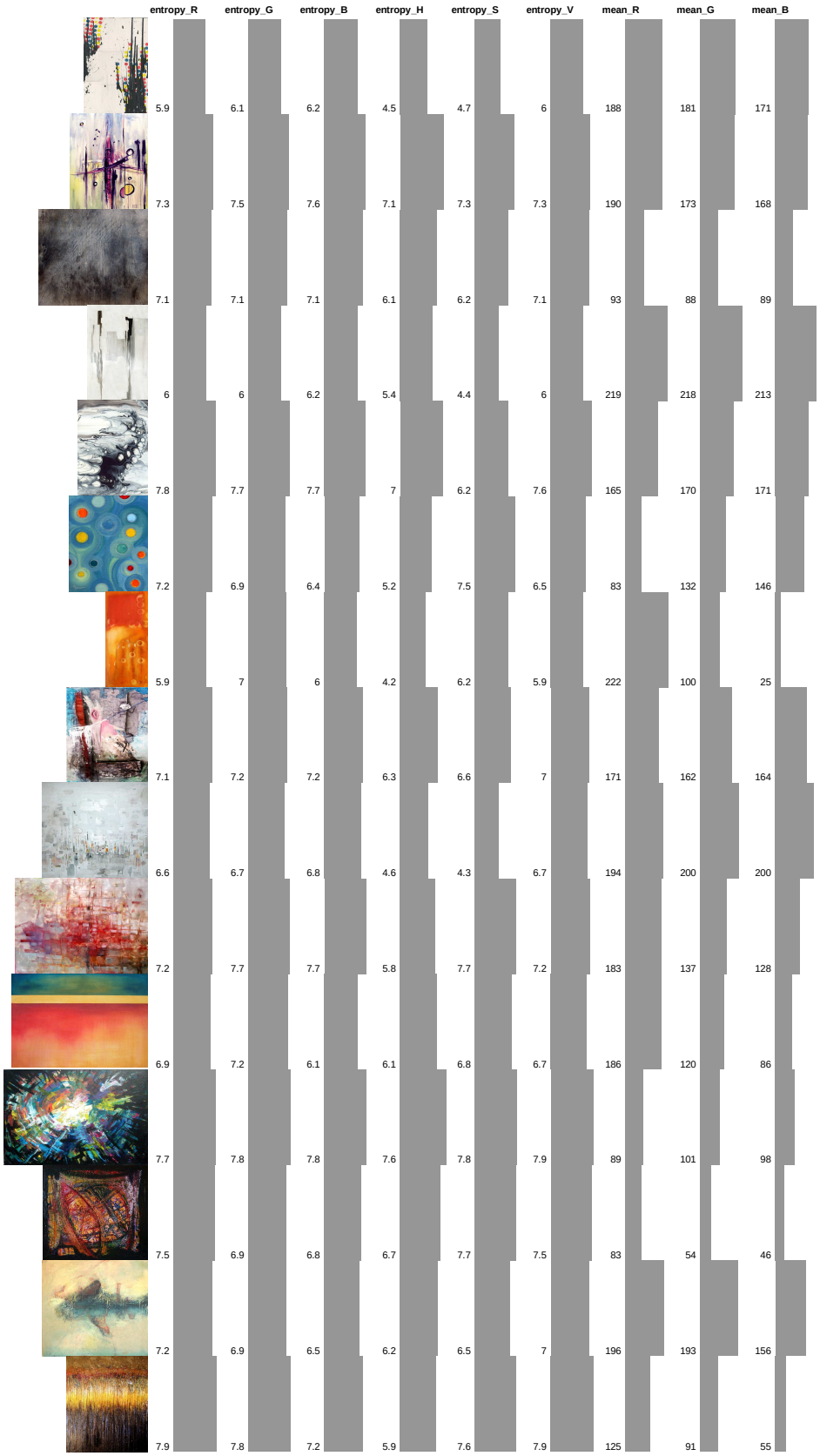


Image features

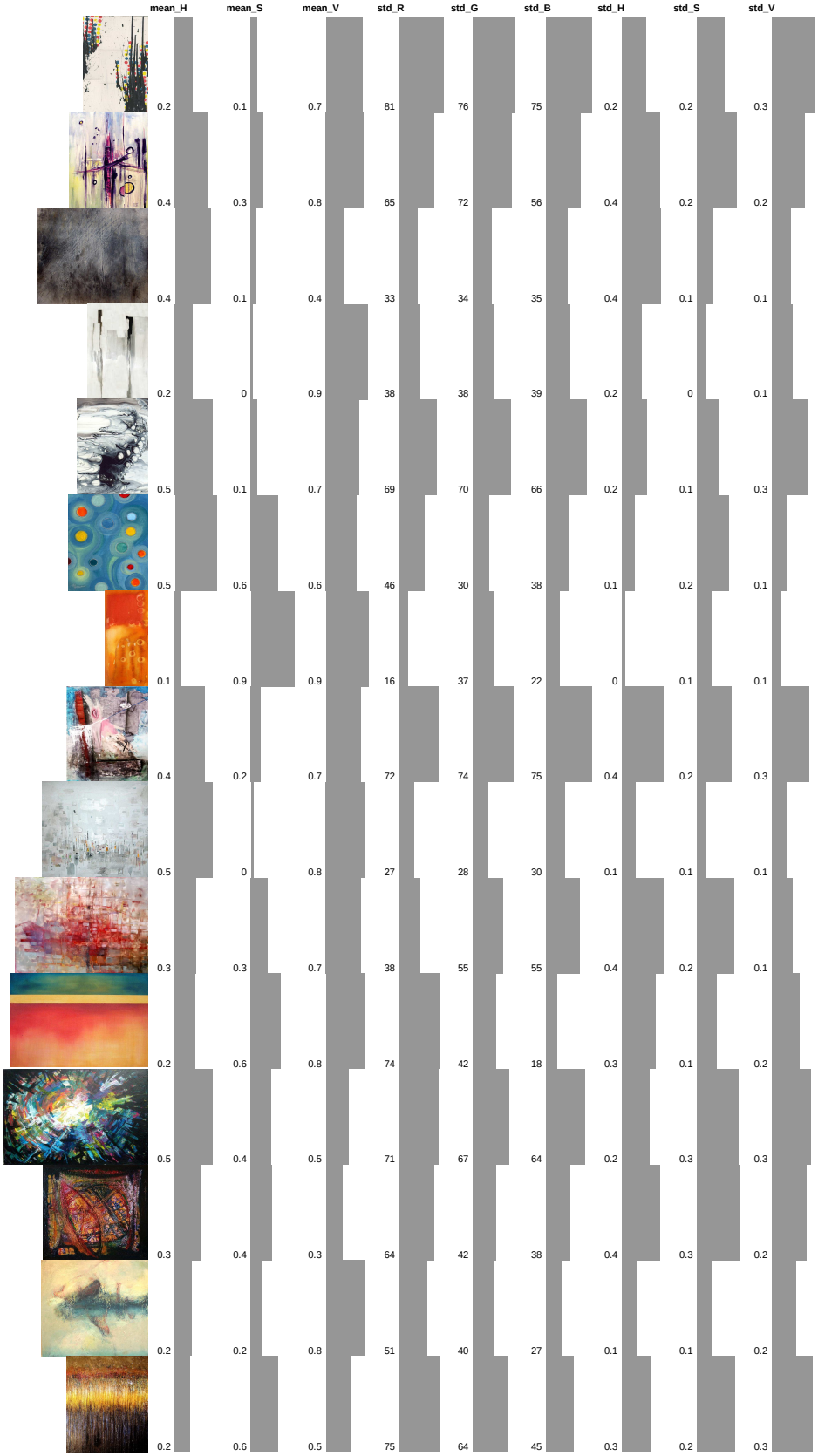
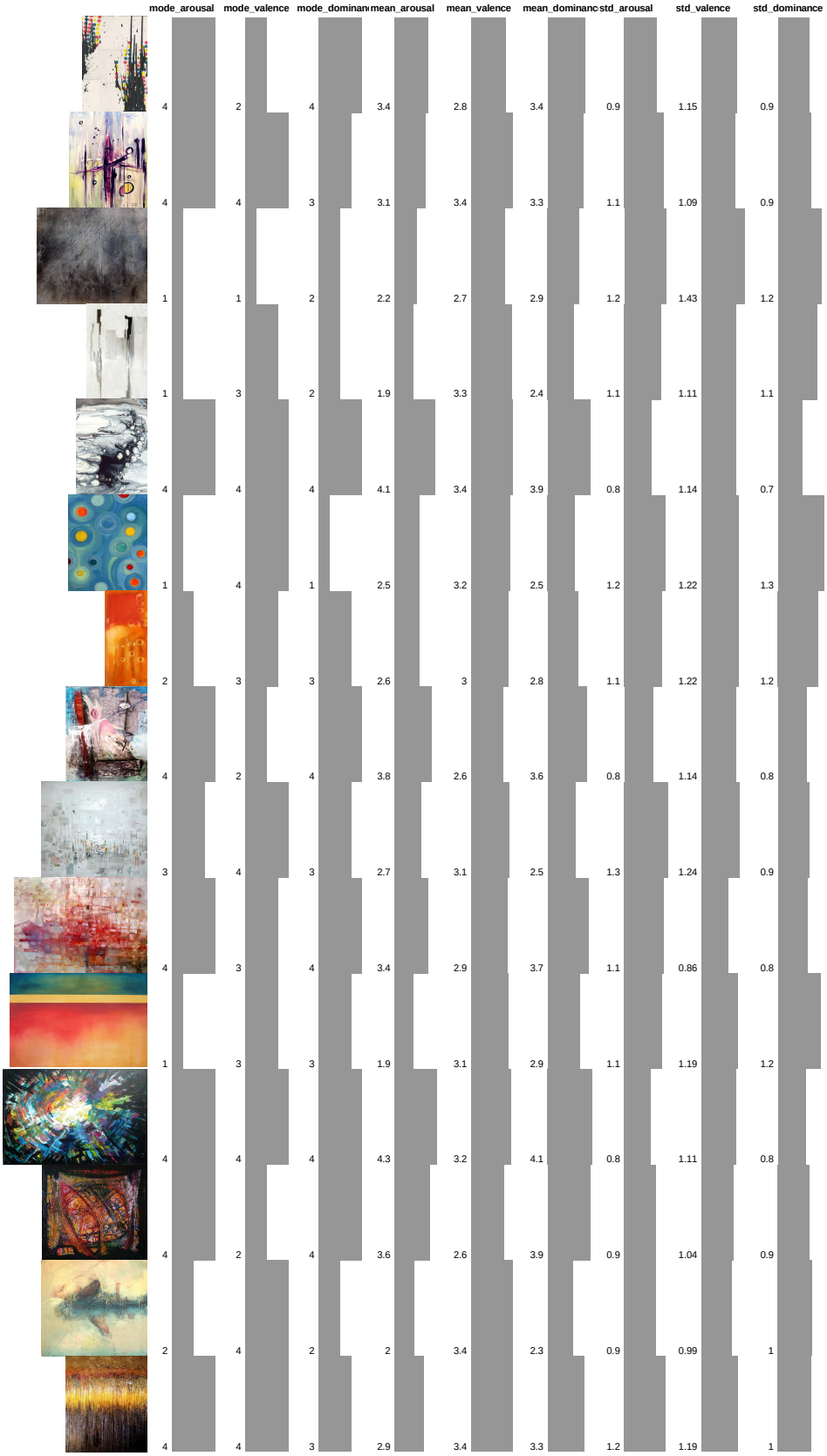


Image features



Correlation analysis on statistics derived from gray scale cooccurence matrices

degrees	0				45			
Hue	Contrast	Correlation	Energy	Homogene	Contrast	Correlation	Energy	Homogene
Mode arousal	**0.0497**	0.3244	0.2002	0.1249	0.0827	0.3104	0.1987	0.1412
Mode valence	0.4815	0.5487	0.4678	0.8455	0.2463	0.751	0.441	0.877
Mode dominance	0.2375	0.8127	0.106	0.3462	0.2439	0.7442	0.1	0.3331
Mean arousal	0.0861	0.3565	0.1011	0.0617	0.1115	0.3153	0.1022	0.0606
Mean valence	0.6882	0.1536	0.7803	0.3924	0.4173	0.2492	0.7199	0.6213
Mean dominance	**0.048**	0.4087	**0.0193**	0.0698	**0.0412**	0.3209	**0.0192**	0.0527
Stdev arousal	0.54	0.5779	0.1854	0.1514	0.6304	0.608	0.2093	0.1862
Stdev valence	0.9321	0.9017	0.7146	0.9209	0.9613	0.7514	0.7043	0.7593
Stdev dominance	0.1757	0.2908	0.1461	0.0858	0.2493	0.2763	0.1432	0.1016

degrees	90				135			
Hue	Contrast	Correlation	Energy	Homogene	Contrast	Correlation	Energy	Homogene
Mode arousal	0.1014	0.4994	0.2395	0.2218	0.0847	0.3152	0.1997	0.1446
Mode valence	0.2916	0.673	0.4584	0.9448	0.2514	0.7396	0.4397	0.8735
Mode dominance	0.2381	0.8365	0.1164	0.3786	0.2492	0.7505	0.1	0.3377
Mean arousal	0.0852	0.3306	0.1103	0.0672	0.1138	0.3205	0.1033	0.0627
Mean valence	0.4947	0.196	0.7615	0.4603	0.429	0.2417	0.7214	0.6197
Mean dominance	**0.0387**	0.3601	**0.0213**	0.0694	**0.042**	0.325	**0.0194**	0.0538
Stdev arousal	0.4249	0.3718	0.1783	0.1067	0.6388	0.6164	0.2091	0.1887
Stdev valence	0.8987	0.8253	0.705	0.8091	0.9488	0.7415	0.7022	0.751
Stdev dominance	0.207	0.3112	0.1588	0.1038	0.2536	0.2809	0.1437	0.1045

degrees	0				45			
Saturation	Contrast	Correlation	Energy	Homogene	Contrast	Correlation	Energy	Homogene
Mode arousal	**0.0234**	0.3018	0.1079	**0.0124**	**0.018**	0.2851	0.1139	**0.0101**
Mode valence	0.4554	0.3553	0.8178	0.6251	0.467	0.3427	0.7953	0.6974
Mode dominance	0.2429	0.9052	0.2879	0.1772	0.2103	0.8756	0.2885	0.1479
Mean arousal	0.0507	0.5248	0.0518	**0.0125**	**0.0311**	0.4537	0.0544	**0.0084**
Mean valence	0.5775	0.1669	0.5727	0.7918	0.6309	0.1994	0.5529	0.9019
Mean dominance	**0.0282**	0.6708	**0.0124**	**0.0063**	**0.0184**	0.6807	**0.0127**	**0.0033**
Stdev arousal	0.4806	0.9642	0.2604	0.2546	0.4044	0.9531	0.2657	0.2393
Stdev valence	0.7932	0.4379	0.4673	0.6728	0.7599	0.3355	0.4731	0.6013
Stdev dominance	0.2096	0.3429	0.5712	0.1541	0.1824	0.2862	0.5786	0.1404

degrees	90				135			
Saturation	Contrast	Correlation	Energy	Homogene	Contrast	Correlation	Energy	Homogene
Mode arousal	**0.0256**	0.2766	0.1177	**0.0173**	**0.0182**	0.2984	0.1155	**0.0107**
Mode valence	0.4005	0.2737	0.8389	0.5618	0.4689	0.3499	0.7938	0.6959
Mode dominance	0.1846	0.7026	0.2929	0.1704	0.2122	0.9014	0.2911	0.1513
Mean arousal	**0.0152**	0.1866	0.0528	**0.0064**	**0.0322**	0.4801	0.0556	**0.0091**
Mean valence	0.6028	0.1617	0.5815	0.7871	0.6274	0.1933	0.5542	0.8911
Mean dominance	**0.0101**	0.3025	**0.0124**	**0.0032**	**0.0188**	0.7044	**0.013**	**0.0035**
Stdev arousal	0.2187	0.3312	0.2545	0.1627	0.407	0.9506	0.2659	0.2388
Stdev valence	0.7736	0.4078	0.4711	0.6564	0.7613	0.3293	0.4747	0.6083
Stdev dominance	0.1467	0.1854	0.5804	0.1427	0.1843	0.299	0.582	0.1434

degrees	0				45			
Value	Contrast	Correlation	Energy	Homogene	Contrast	Correlation	Energy	Homogene
Mode arousal	**0.0113**	0.5708	**0.0264**	**0.0282**	**0.003**	0.6615	**0.0241**	**0.0097**
Mode valence	0.793	0.7431	0.9923	0.8972	0.9505	0.8856	0.9375	0.9485
Mode dominance	0.1469	0.1279	0.1612	0.3232	0.064	0.1836	0.1463	0.1744
Mean arousal	0.0895	0.4808	**0.0376**	0.0759	**0.021**	0.7067	**0.0321**	**0.0201**
Mean valence	0.863	0.7528	0.9746	0.9881	0.9433	0.9092	0.9589	0.9405
Mean dominance	0.0653	0.2448	**0.0126**	**0.0403**	**0.0184**	0.2943	**0.0108**	**0.0103**
Stdev arousal	0.4933	**0.024**	0.1377	0.471	0.232	**0.0414**	0.1339	0.2656
Stdev valence	0.9169	**0.0414**	0.4741	0.9986	0.8221	0.0608	0.4297	0.8225
Stdev dominance	0.1397	0.2013	**0.037**	0.2011	**0.0438**	0.35	**0.03**	0.0763

degrees	90				135			
Value	Contrast	Correlation	Energy	Homogene	Contrast	Correlation	Energy	Homogene
Mode arousal	**0.0064**	0.2566	**0.0338**	**0.0359**	**0.0038**	0.5898	**0.0246**	**0.0113**
Mode valence	0.702	0.6604	0.8467	0.7078	0.9046	0.9939	0.9564	0.9943
Mode dominance	0.0706	0.0763	0.181	0.2883	0.0773	0.1554	0.1464	0.1872
Mean arousal	**0.0033**	0.6787	**0.0334**	**0.0177**	**0.0306**	0.5951	**0.0338**	**0.0252**
Mean valence	0.9124	0.952	0.9227	0.8563	0.9271	0.8659	0.9698	0.9691
Mean dominance	**0.0033**	0.2942	**0.0121**	**0.0117**	**0.0256**	0.2467	**0.0111**	**0.0123**
Stdev arousal	0.0607	0.1636	0.1252	0.1744	0.2855	**0.0287**	0.1376	0.2943
Stdev valence	0.7762	0.1061	0.4559	0.8896	0.8481	**0.0478**	0.4405	0.8695
Stdev dominance	**0.0227**	0.1425	**0.0408**	0.1183	0.0561	0.2886	**0.0307**	0.087

Correlation analysis on statistics derived from edge detection algortihms

	Sobel edge on Hue		
Direction	**Both**	**Horizontal**	**Vertical**
Mode arousal	0.1098	0.1947	0.0789
Mode valence	0.1745	0.5775	0.5996
Mode dominance	0.3543	0.48	0.3304
Mean arousal	0.2219	0.2442	0.1548
Mean valence	0.5449	0.7549	0.8118
Mean dominance	0.1189	0.1419	0.1282
Stdev arousal	0.6857	0.5372	0.4593
Stdev valence	0.8112	0.9478	0.9111
Stdev dominance	0.2777	0.2441	0.1524
	Sobel edge on Saturation		
Direction	**Both**	**Horizontal**	**Vertical**
Mode arousal	**0.0171**	0.0774	0.0771
Mode valence	0.7983	0.8773	0.8137
Mode dominance	0.0993	0.1506	0.1211
Mean arousal	**0.0134**	**0.0281**	0.0689
Mean valence	0.7314	0.8713	0.7975
Mean dominance	**0.0033**	**0.0047**	**0.005**
Stdev arousal	**0.0194**	**0.0346**	0.0598
Stdev valence	0.2097	0.3451	0.3175
Stdev dominance	**0.0376**	0.1142	0.1637
	Sobel edge on Value		
Direction	**Both**	**Horizontal**	**Vertical**
Mode arousal	**0.0007**	**0.0327**	0.0549
Mode valence	0.8658	0.7763	0.4988
Mode dominance	0.0858	0.2759	0.075
Mean arousal	**0.0038**	**0.0233**	0.1559
Mean valence	0.7314	0.5298	0.427
Mean dominance	**0.0138**	**0.0262**	**0.0131**
Stdev arousal	0.1543	0.3201	0.5142
Stdev valence	0.3851	0.6415	0.7708
Stdev dominance	**0.0113**	0.0915	0.3978
	Canny edge		
	Hue	**Saturation**	**Value**
Mode arousal	0.3769	0.4267	0.6813
Mode valence	0.445	0.818	0.5027
Mode dominance	0.6929	0.7545	0.4788
Mean arousal	0.1764	0.3948	0.6188
Mean valence	0.1361	0.5562	0.6958
Mean dominance	0.3636	0.133	0.8671
Stdev arousal	0.1023	0.2798	0.3018
Stdev valence	0.9119	0.7555	0.2256
Stdev dominance	0.1559	0.7043	0.3758

Correlation analysis on statistics derived from color histograms

	Red									
Bin	**0**	**28**	**56**	**85**	**113**	**141**	**170**	**198**	**226**	**255**
Mode arousal	0.1305	0.7186	0.1898	0.1116	0.2978	0.9812	0.2071	0.4798	0.1674	0.4996
Mode valence	0.6374	0.3739	**0.0488**	**0.0393**	**0.0332**	0.1147	0.6009	0.2600	0.6579	0.8596
Mode dominance	0.2230	0.5559	0.1590	0.1500	0.3248	0.8347	0.2774	0.5360	0.8073	0.5046
Mean arousal	0.0919	0.3385	0.6802	0.3930	0.6385	0.9180	0.6031	0.8718	**0.0414**	0.9110
Mean valence	0.9139	0.2202	0.1546	0.2206	0.2005	0.2663	0.8451	0.7194	0.4205	0.9328
Mean dominance	**0.0286**	0.0890	0.8751	0.7113	0.9718	0.5490	0.6118	0.6555	0.1664	0.8983
Stdev arousal	0.2788	0.4147	0.4528	0.3842	0.4830	0.4340	0.3315	0.2504	0.6322	0.1152
Stdev valence	0.7014	0.8245	**0.0139**	**0.0153**	**0.0388**	0.5161	0.0970	0.1196	0.5297	0.3651
Stdev dominance	0.3530	0.9020	0.0762	0.0861	0.2539	0.9873	0.1119	0.2352	0.3976	0.4215

	Green									
Bin	**0**	**28**	**56**	**85**	**113**	**141**	**170**	**198**	**226**	**255**
Mode arousal	**0.0476**	0.2691	0.3446	0.1787	**0.0287**	**0.0434**	0.8556	0.7638	0.9429	0.2001
Mode valence	0.7256	0.1377	**0.0023**	**0.0166**	0.1517	0.4538	0.4043	0.0804	0.5321	0.9955
Mode dominance	0.3393	0.2450	0.9447	0.5386	0.1148	0.1293	0.8355	0.8142	0.7445	0.2379
Mean arousal	0.1803	0.1895	0.5553	0.3510	0.2192	0.2722	0.6456	0.5142	0.4812	0.1099
Mean valence	0.8021	0.0890	**0.0391**	0.1356	0.3035	0.4646	0.9478	0.2077	0.2625	0.8235
Mean dominance	0.0685	**0.0402**	0.6514	0.9736	0.6276	0.6049	0.8110	0.2200	0.1616	0.3899
Stdev arousal	0.8641	0.4157	0.4437	0.2881	0.1952	0.1865	0.2506	0.5998	0.8599	**0.0462**
Stdev valence	0.6989	0.7995	0.0631	0.0610	0.0511	0.2218	0.2533	0.1556	0.4995	0.7815
Stdev dominance	0.3514	0.5174	0.1973	0.1215	**0.0096**	**0.0248**	0.9558	0.3001	0.4192	0.0520

	Blue									
Bin	**0**	**28**	**56**	**85**	**113**	**141**	**170**	**198**	**226**	**255**
Mode arousal	0.4947	0.2980	0.3897	0.1049	0.1487	0.1896	0.7362	0.3521	0.7243	0.2314
Mode valence	0.8274	0.2022	**0.0109**	0.1113	0.0788	0.9300	0.1387	0.2151	0.9184	0.7373
Mode dominance	0.5929	0.2749	0.9797	0.7230	0.5647	**0.0191**	0.1899	0.6347	0.7462	0.2148
Mean arousal	0.6249	0.3213	0.5294	0.1933	0.2953	0.4088	0.4779	0.7258	0.9911	0.1057
Mean valence	0.5703	0.1476	0.0923	0.3097	0.3124	0.8452	0.2706	0.6004	0.6827	0.4046
Mean dominance	0.4310	0.0878	0.5951	0.9328	0.9890	0.2698	0.1699	0.6463	0.5495	0.3729
Stdev arousal	0.8693	0.7815	0.4356	0.4624	0.4350	0.3714	0.8891	0.4390	0.9109	0.1099
Stdev valence	0.9460	0.9846	0.0894	0.1751	0.2037	0.8369	0.1013	0.4919	0.9418	0.9705
Stdev dominance	0.8202	0.7637	0.2327	0.1051	0.1778	0.1978	0.8872	0.1135	0.3768	0.0964

	Hue									
Bin	**0.00**	**0.11**	**0.22**	**0.33**	**0.44**	**0.55**	**0.66**	**0.77**	**0.88**	**1.00**
Mode arousal	0.9292	0.4631	0.2112	0.4029	0.8441	0.4838	0.4330	0.6668	0.8575	0.8949
Mode valence	0.0527	0.3066	0.6322	0.2490	0.1432	0.2337	**0.0428**	0.1245	**0.0268**	**0.0262**
Mode dominance	0.3314	0.5575	0.2261	0.8554	0.7922	0.2591	0.5332	0.5442	0.9141	0.3164
Mean arousal	0.9990	0.0694	0.1802	0.5362	0.9842	0.9737	0.8176	0.8608	0.6769	0.8575
Mean valence	**0.0358**	0.9941	0.1037	0.2326	0.5825	0.8125	0.3051	0.5756	0.1357	**0.0216**
Mean dominance	0.2808	0.4268	0.1808	0.4621	0.4330	0.3541	0.8253	0.7889	0.3480	0.2115
Stdev arousal	0.5095	0.5122	0.3761	0.1774	0.8993	0.3831	0.6110	0.4492	0.6802	0.6844
Stdev valence	0.4513	0.9859	0.4751	0.8025	0.5518	0.2024	**0.0222**	0.0666	0.3202	0.5776
Stdev dominance	0.8350	0.3068	0.9148	0.6217	0.6110	0.6317	0.4566	0.5227	0.9240	0.9152

	Saturation									
Bin	**0.00**	**0.11**	**0.22**	**0.33**	**0.44**	**0.55**	**0.66**	**0.77**	**0.88**	**1.00**
Mode arousal	0.4354	0.9932	0.8713	0.9434	0.4525	0.4117	0.2594	0.9885	0.9458	0.7089
Mode valence	0.4318	**0.0110**	0.5917	0.8975	0.6321	0.7801	0.8573	0.6270	0.9623	0.8244
Mode dominance	0.5359	0.9234	0.5189	0.6924	0.7519	0.9047	0.8323	0.8722	0.6800	0.8066
Mean arousal	0.4671	0.8902	0.6802	0.7539	0.3604	0.5455	0.2995	0.8163	0.9027	0.8925
Mean valence	0.7215	0.1010	0.9122	0.9454	0.8188	0.8779	0.9593	0.9277	0.7678	0.9211
Mean dominance	0.2835	0.7540	0.8325	0.9067	0.1470	0.7768	0.9193	0.5351	0.9375	0.7689
Stdev arousal	0.2481	0.7804	0.7410	0.5753	0.9534	0.5745	0.5506	0.4374	0.7037	0.6447
Stdev valence	0.0759	0.1833	0.3734	0.0659	0.0781	0.7942	0.8822	0.8676	0.7958	0.6094
Stdev dominance	0.9877	0.9100	0.8928	0.9594	0.8990	0.1985	0.1416	0.3654	0.4742	0.2898

	Value									
Bin	**0.00**	**0.11**	**0.22**	**0.33**	**0.44**	**0.55**	**0.66**	**0.77**	**0.88**	**1.00**
Mode arousal	0.1340	0.2690	0.7485	0.2855	0.1914	0.2292	0.9438	0.5019	0.1976	0.4394
Mode valence	0.7343	0.2428	**0.0039**	**0.0138**	**0.0258**	0.4811	0.3158	0.2027	0.5435	0.9509
Mode dominance	0.4784	0.2583	0.9065	0.4497	0.4614	0.0838	0.5371	0.6089	0.8169	0.4597
Mean arousal	0.3752	0.1738	0.9264	0.5455	0.4050	0.7146	0.8167	0.8978	**0.0427**	0.8023
Mean valence	0.7763	0.1668	0.0685	0.1603	0.1985	0.5027	0.9809	0.6378	0.4078	0.8889
Mean dominance	0.1835	**0.0440**	0.5276	0.9871	0.9755	0.8387	0.8911	0.6702	0.1266	0.9183
Stdev arousal	0.9812	0.3092	0.8166	0.4969	0.4920	0.2291	0.1939	0.3630	0.5295	0.0864
Stdev valence	0.9440	0.8728	**0.0386**	**0.0257**	**0.0393**	0.2361	0.3136	0.0528	0.5918	0.3731
Stdev dominance	0.5074	0.5016	0.5850	0.2640	0.1804	0.1048	0.9620	0.2383	0.4798	0.2879

Feature selection on the emotion dimensions

Arousal

Feature	Avg. Rank	#
edge_V_Sobel_Both	4.70	**1**
std_G	5.80	**2**
std_B	7.33	**3**
std_S	7.85	**4**
SatEnergy0	9.05	**5**
SatEnergy90	12.73	6
SatHomogeneity45	13.18	7
edge_S_Sobel_Both	13.28	8
SatEnergy45	18.83	9
entropy_B	20.70	**10**
HueContrast135	20.73	11
SatHomogeneity0	21.28	12
HueContrast90	22.70	13
std_V	23.43	14
SatContrast45	23.45	15
SatHomogeneity90	24.20	16
std_H	24.60	17
ValContrast135	24.95	18
HueContrast45	26.10	19
HueContrast0	26.50	**20**
std_R	26.68	21
HueEnergy90	27.35	22
ValContrast45	27.93	23
SatEnergy135	28.08	24
HueEnergy135	28.28	25
SatContrast0	29.00	26
SatContrast135	29.00	27
SatContrast90	29.00	28
edge_S_Sobel_Hori	29.28	29
SatHomogeneity135	29.48	30
entropy_V	29.50	31
HueEnergy0	29.60	32
entropy_S	30.25	33
ValEnergy135	30.30	34
HueEnergy45	30.73	35
entropy_H	31.18	36
entropy_G	31.58	37
edge_V_Sobel_Hori	33.15	38
ValCorrelation135	33.20	39
ValHomogeneity135	34.10	40
ValEnergy45	34.25	41
ValContrast90	34.55	42
entropy_R	35.13	43
edge_S_Sobel_Vert	35.35	44
ValEnergy0	35.50	45
edge_V_Sobel_Vert	37.63	46
ValContrast0	37.80	47
ValHomogeneity90	38.45	48
ValEnergy90	38.80	49
ValCorrelation45	39.00	50
ValCorrelation0	39.93	51
ValHomogeneity45	40.88	52
ValCorrelation90	44.05	53
ValHomogeneity0	44.53	54

Dominance

Feature	Avg. Rank
std_S	6.40
edge_V_Sobel_Both	10.35
std_G	12.03
HueEnergy135	14.13
std_B	15.10
SatHomogeneity45	17.25
SatContrast45	18.48
SatEnergy90	18.93
SatEnergy0	20.40
SatContrast0	21.08
HueEnergy45	22.53
SatHomogeneity90	22.73
SatEnergy45	23.05
SatContrast135	23.40
SatHomogeneity0	23.40
entropy_G	23.78
HueEnergy0	24.00
HueEnergy90	24.35
SatEnergy135	24.43
std_R	24.53
entropy_H	25.38
entropy_R	25.70
std_V	26.28
SatContrast90	26.45
entropy_B	27.55
entropy_S	27.70
edge_V_Sobel_Hori	27.88
ValCorrelation90	28.33
HueContrast0	28.35
HueContrast90	29.50
SatHomogeneity135	29.90
ValCorrelation135	30.25
edge_V_Sobel_Vert	30.45
std_H	30.63
HueContrast135	30.85
entropy_V	31.18
ValContrast90	31.30
ValCorrelation0	31.80
ValContrast135	32.00
ValContrast45	32.25
ValEnergy135	32.28
ValCorrelation45	32.88
edge_S_Sobel_Hori	33.08
ValHomogeneity135	33.60
edge_S_Sobel_Both	34.10
HueContrast45	35.18
ValContrast0	35.20
ValEnergy0	35.50
ValEnergy45	36.93
edge_S_Sobel_Vert	39.15
ValEnergy90	39.23
ValHomogeneity0	39.95
ValHomogeneity45	41.68
ValHomogeneity90	42.23

Valence

Feature	Avg. Rank	#
ValContrast90	4.40	**1**
ValHomogeneity90	13.40	**2**
entropy_R	14.50	**3**
SatContrast0	17.00	**4**
SatContrast45	17.68	**5**
ValHomogeneity135	18.00	6
SatContrast135	18.73	7
SatHomogeneity0	19.28	8
entropy_V	19.63	9
entropy_B	19.75	**10**
SatContrast90	20.23	11
ValHomogeneity0	20.33	12
std_R	21.35	13
std_V	21.43	14
entropy_G	21.58	15
ValHomogeneity45	21.60	16
SatHomogeneity90	22.15	17
HueContrast135	22.50	18
SatEnergy45	22.53	19
std_B	24.38	**20**
SatHomogeneity45	24.40	21
ValContrast135	24.73	22
std_H	24.95	23
edge_V_Sobel_Hori	25.40	24
HueContrast45	25.63	25
SatEnergy0	26.48	26
std_G	26.60	27
ValContrast45	27.05	28
HueEnergy135	27.25	29
std_S	28.23	30
HueContrast90	28.28	31
SatEnergy90	28.55	32
HueEnergy90	28.78	33
edge_V_Sobel_Both	28.93	34
entropy_H	30.00	35
HueContrast0	30.93	36
entropy_S	31.45	37
HueEnergy45	31.65	38
ValEnergy0	33.15	39
ValContrast0	33.20	40
HueEnergy0	33.23	41
SatHomogeneity135	35.30	42
ValEnergy90	35.55	43
ValCorrelation0	35.93	44
edge_V_Sobel_Vert	36.18	45
ValEnergy135	36.83	46
SatEnergy135	37.03	47
edge_S_Sobel_Both	37.85	48
ValEnergy45	39.88	49
ValCorrelation135	41.38	50
ValCorrelation45	42.00	51
ValCorrelation90	42.13	52
edge_S_Sobel_Hori	42.35	53
edge_S_Sobel_Vert	43.40	54

All

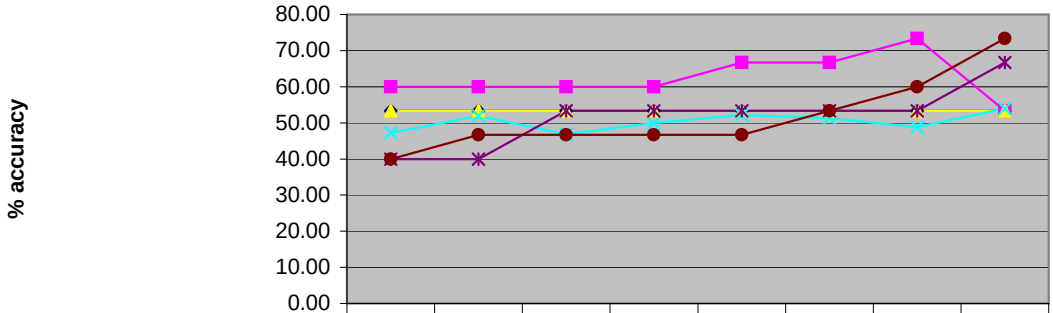
Feature	Avg. Rank
std_S	14.16
edge_V_Sobel_Both	14.66
std_G	14.81
std_B	15.60
SatHomogeneity45	18.28
SatEnergy0	18.64
SatContrast45	19.87
SatEnergy90	20.07
SatHomogeneity0	21.32
SatEnergy45	21.47
SatContrast0	22.36
entropy_B	22.67
SatHomogeneity90	23.03
HueEnergy135	23.22
ValContrast90	23.42
SatContrast135	23.71
std_V	23.71
std_R	24.18
HueContrast135	24.69
entropy_R	25.11
SatContrast90	25.23
entropy_G	25.64
std_H	26.73
entropy_V	26.77
HueContrast90	26.83
HueEnergy90	26.83
ValContrast135	27.23
HueEnergy45	28.30
edge_S_Sobel_Both	28.41
ValHomogeneity135	28.57
HueContrast0	28.59
edge_V_Sobel_Hori	28.81
entropy_H	28.85
HueEnergy0	28.94
HueContrast45	28.97
ValContrast45	29.08
entropy_S	29.80
SatEnergy135	29.84
ValHomogeneity90	31.36
SatHomogeneity135	31.56
ValEnergy135	33.13
ValEnergy0	34.72
ValHomogeneity45	34.72
edge_V_Sobel_Vert	34.75
edge_S_Sobel_Hori	34.90
ValHomogeneity0	34.93
ValCorrelation135	34.94
ValContrast0	35.40
ValCorrelation0	35.88
ValEnergy45	37.02
ValEnergy90	37.86
ValCorrelation45	37.96
ValCorrelation90	38.17
edge_S_Sobel_Vert	39.30

WEKA classifiers and configuration

classifier	configuration parameters
bayes.BayesNet	'-Q bayes.net.search.local.K2 -- -P 1 -S BAYES -E bayes.net.estimate.BMAEstimator -- -A 0.5' 746037443258775954
bayes.NaiveBayes	'' 5995231201785697655
functions.MultilayerPerceptron	'-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H o -D' 572250905027665169
trees.RandomForest	'-I 10 -K 0 -S 1' 4216839470751428698
trees.NBTree	'' -4716005707058256086
meta.AdaBoostM1	'-P 100 -S 1 -I 10 -W trees.DecisionStump' -7378107808933117974

Machine learning results

Accuracy on arousal

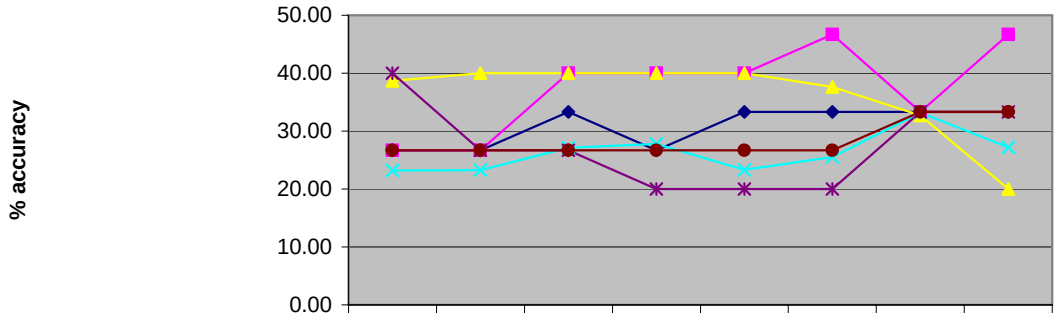


	114	20	10	5	4	3	2	1
Bayes Net	53.33	53.33	53.33	53.33	53.33	53.33	53.33	53.33
Naive Bayes	60.00	60.00	60.00	60.00	66.67	66.67	73.33	53.33
Multilayer Perceptron	53.33	53.33	53.33	53.33	53.33	53.33	53.33	53.33
Random Forest	47.20	51.93	46.87	50.00	52.20	51.27	48.80	53.93
Naïve Bayes Tree	40.00	40.00	53.33	53.33	53.33	53.33	53.33	66.67
Decision Stump with Ada Boost	40.00	46.67	46.67	46.67	46.67	53.33	60.00	73.33

% accuracy

number of features

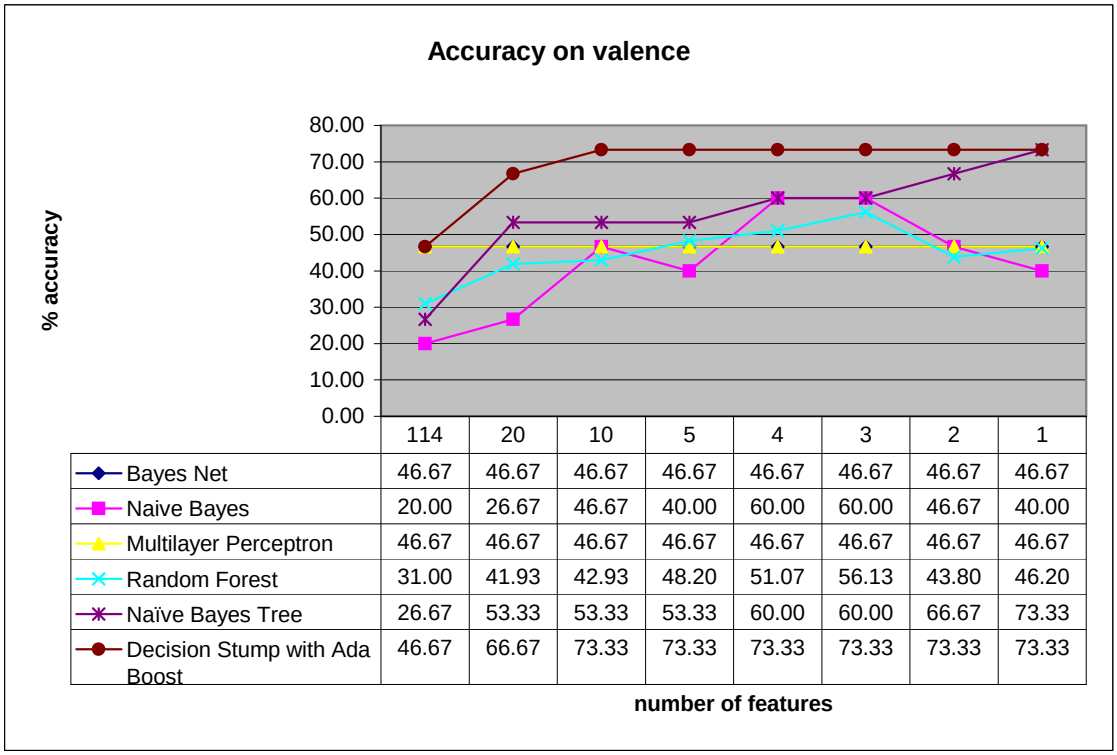
Accuracy on dominance



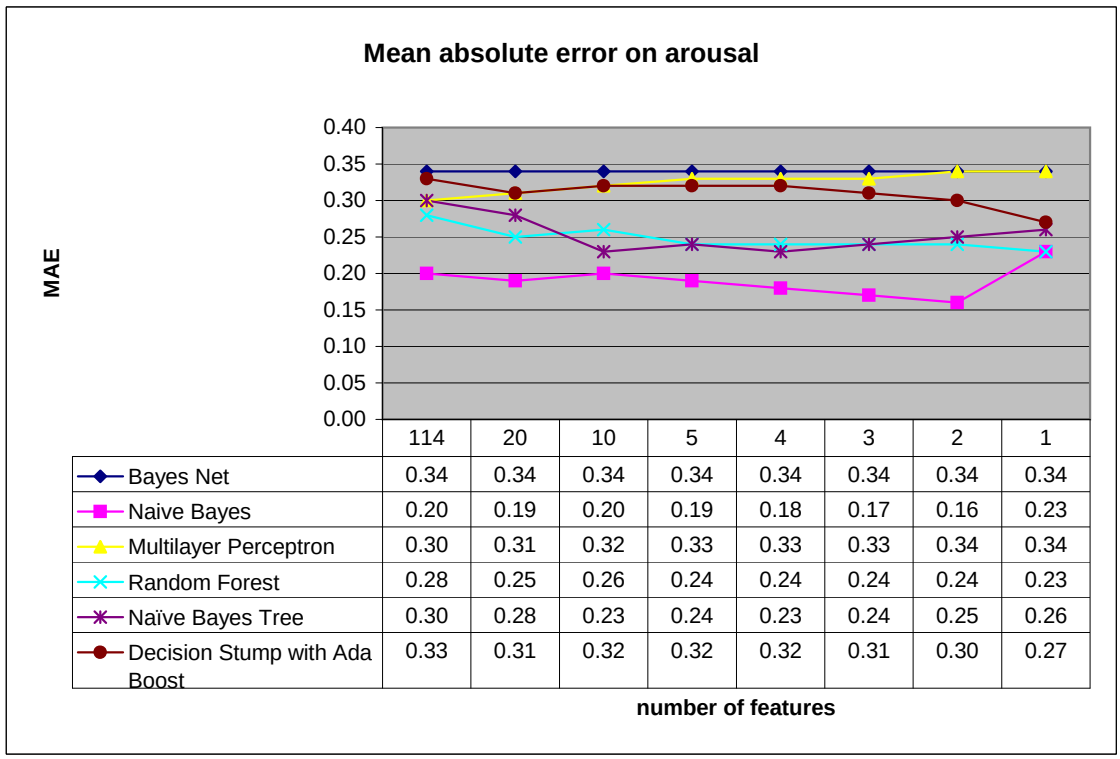
	114	20	10	5	4	3	2	1
Bayes Net	26.67	26.67	33.33	26.67	33.33	33.33	33.33	33.33
Naive Bayes	26.67	26.67	40.00	40.00	40.00	46.67	33.33	46.67
Multilayer Perceptron	38.67	40.00	40.00	40.00	40.00	37.60	32.73	20.00
Random Forest	23.20	23.27	27.13	27.80	23.33	25.53	33.20	27.20
Naïve Bayes Tree	40.00	26.67	26.67	20.00	20.00	20.00	33.33	33.33
Decision Stump with Ada Boost	26.67	26.67	26.67	26.67	26.67	26.67	33.33	33.33

% accuracy

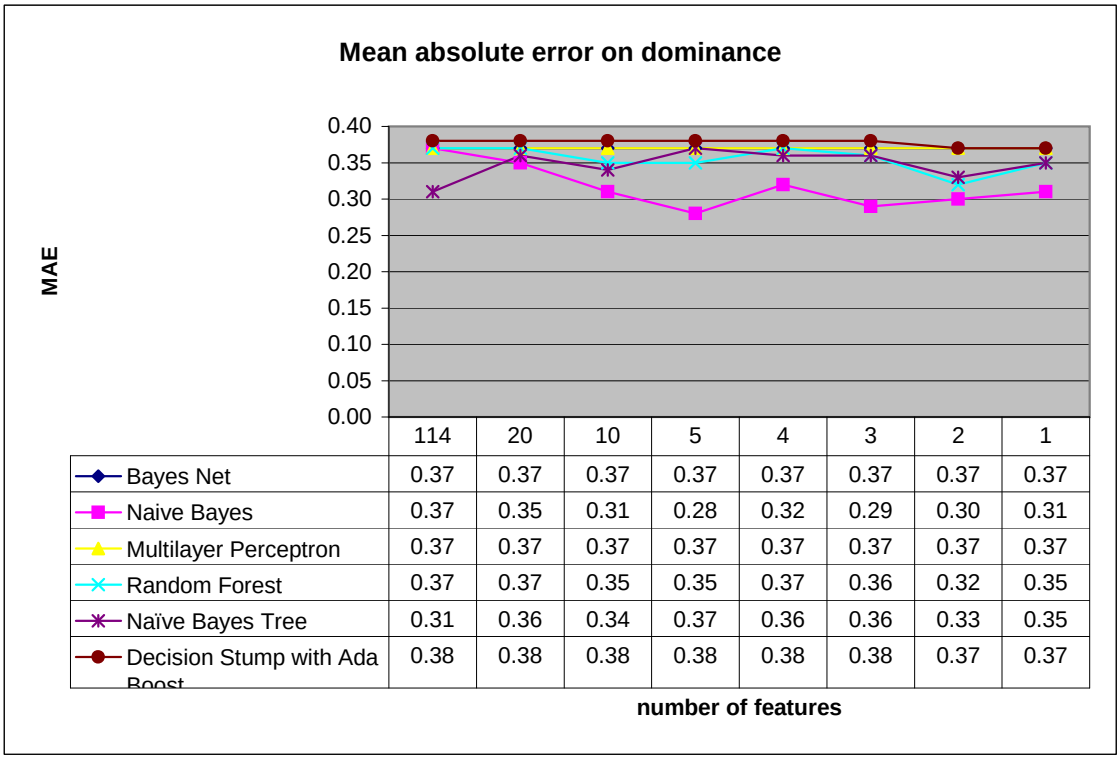
number of features



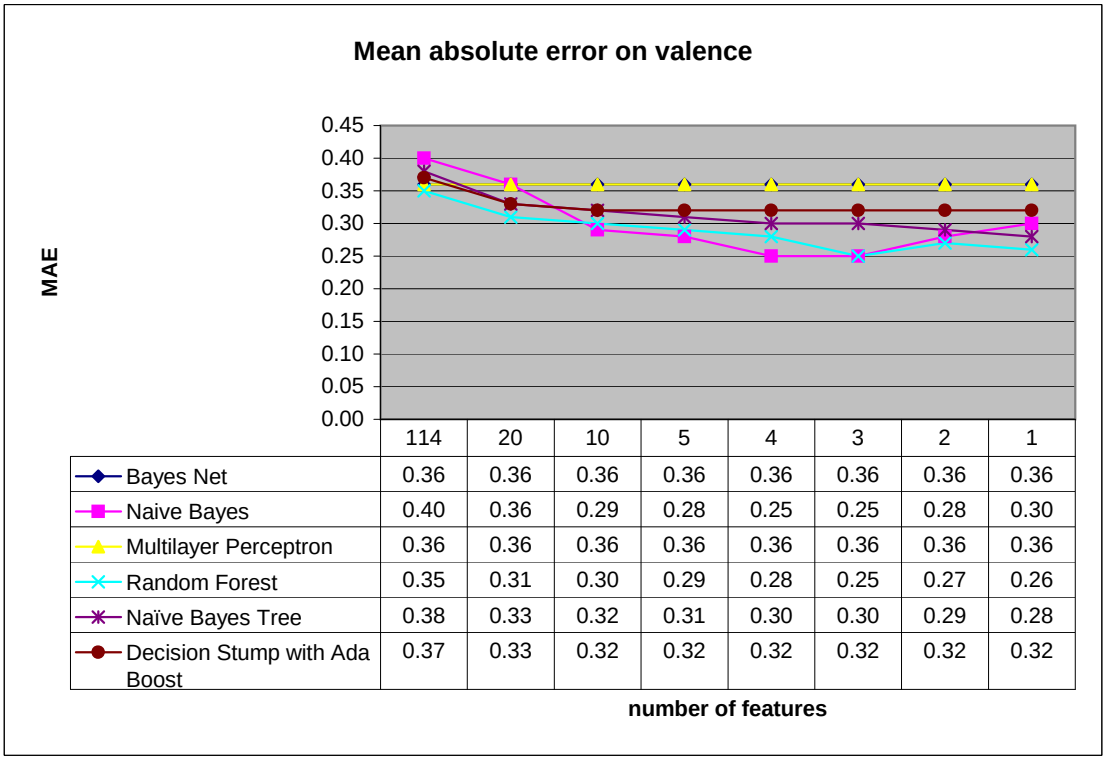
	114	20	10	5	4	3	2	1
Bayes Net	46.67	46.67	46.67	46.67	46.67	46.67	46.67	46.67
Naive Bayes	20.00	26.67	46.67	40.00	60.00	60.00	46.67	40.00
Multilayer Perceptron	46.67	46.67	46.67	46.67	46.67	46.67	46.67	46.67
Random Forest	31.00	41.93	42.93	48.20	51.07	56.13	43.80	46.20
Naïve Bayes Tree	26.67	53.33	53.33	53.33	60.00	60.00	66.67	73.33
Decision Stump with Ada Boost	46.67	66.67	73.33	73.33	73.33	73.33	73.33	73.33



	114	20	10	5	4	3	2	1
Bayes Net	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34
Naive Bayes	0.20	0.19	0.20	0.19	0.18	0.17	0.16	0.23
Multilayer Perceptron	0.30	0.31	0.32	0.33	0.33	0.33	0.34	0.34
Random Forest	0.28	0.25	0.26	0.24	0.24	0.24	0.24	0.23
Naïve Bayes Tree	0.30	0.28	0.23	0.24	0.23	0.24	0.25	0.26
Decision Stump with Ada Boost	0.33	0.31	0.32	0.32	0.32	0.31	0.30	0.27

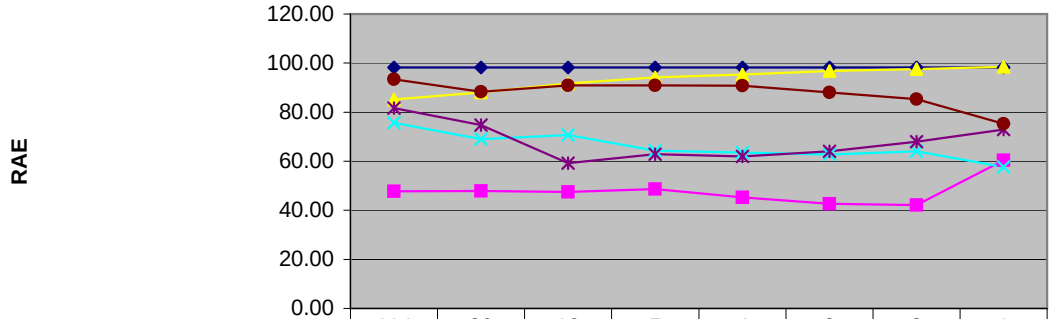


	114	20	10	5	4	3	2	1
Bayes Net	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Naive Bayes	0.37	0.35	0.31	0.28	0.32	0.29	0.30	0.31
Multilayer Perceptron	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Random Forest	0.37	0.37	0.35	0.35	0.37	0.36	0.32	0.35
Naïve Bayes Tree	0.31	0.36	0.34	0.37	0.36	0.36	0.33	0.35
Decision Stump with Ada Boost	0.38	0.38	0.38	0.38	0.38	0.38	0.37	0.37



	114	20	10	5	4	3	2	1
Bayes Net	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36
Naive Bayes	0.40	0.36	0.29	0.28	0.25	0.25	0.28	0.30
Multilayer Perceptron	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36
Random Forest	0.35	0.31	0.30	0.29	0.28	0.25	0.27	0.26
Naïve Bayes Tree	0.38	0.33	0.32	0.31	0.30	0.30	0.29	0.28
Decision Stump with Ada Boost	0.37	0.33	0.32	0.32	0.32	0.32	0.32	0.32

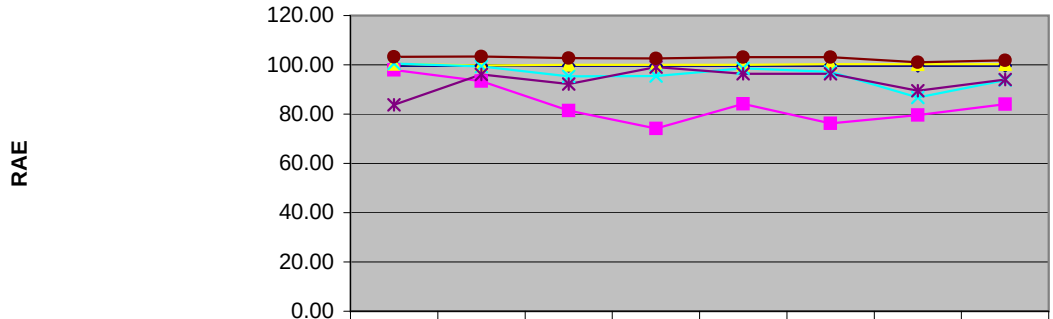
Relative absolute error on arousal



	114	20	10	5	4	3	2	1
Bayes Net	98.22	98.20	98.19	98.19	98.19	98.19	98.19	98.19
Naive Bayes	47.77	47.81	47.47	48.69	45.21	42.67	42.13	60.41
Multilayer Perceptron	85.17	87.99	91.66	94.12	95.35	96.80	97.53	98.52
Random Forest	75.61	68.96	70.66	64.35	63.54	62.78	64.09	57.62
Naïve Bayes Tree	81.63	74.71	59.26	62.83	61.94	64.04	67.90	72.85
Decision Stump with Ada Boost	93.39	88.33	90.91	90.88	90.82	88.10	85.36	75.26

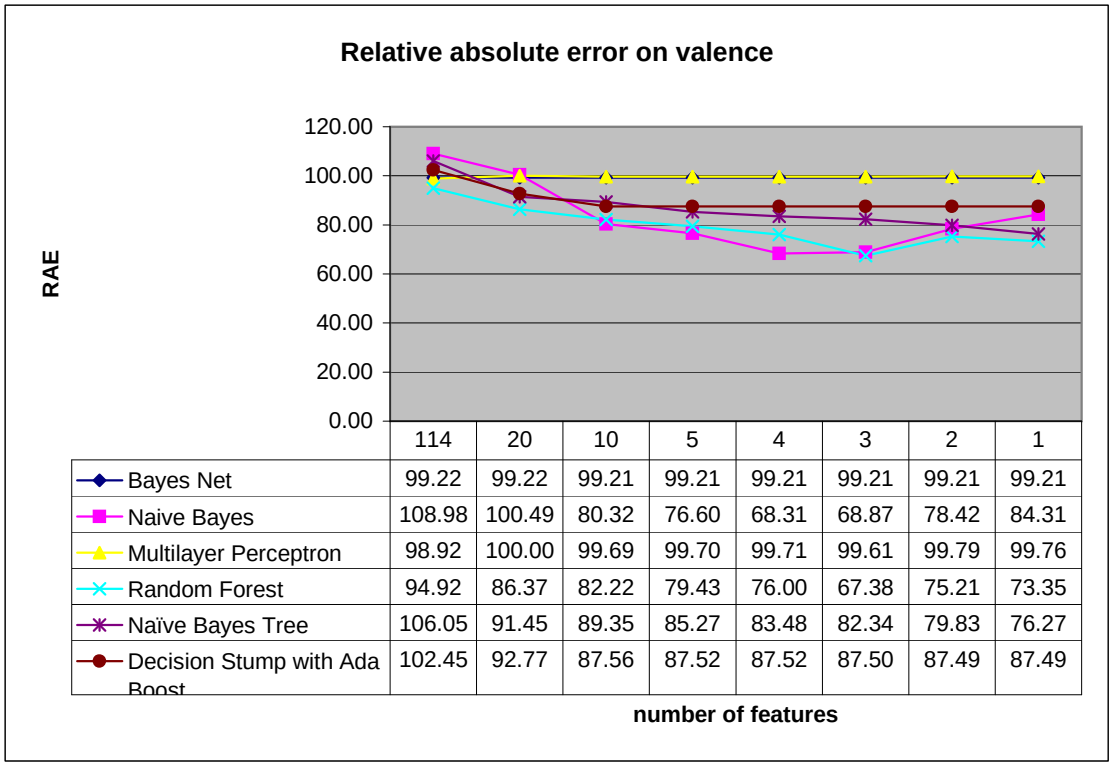
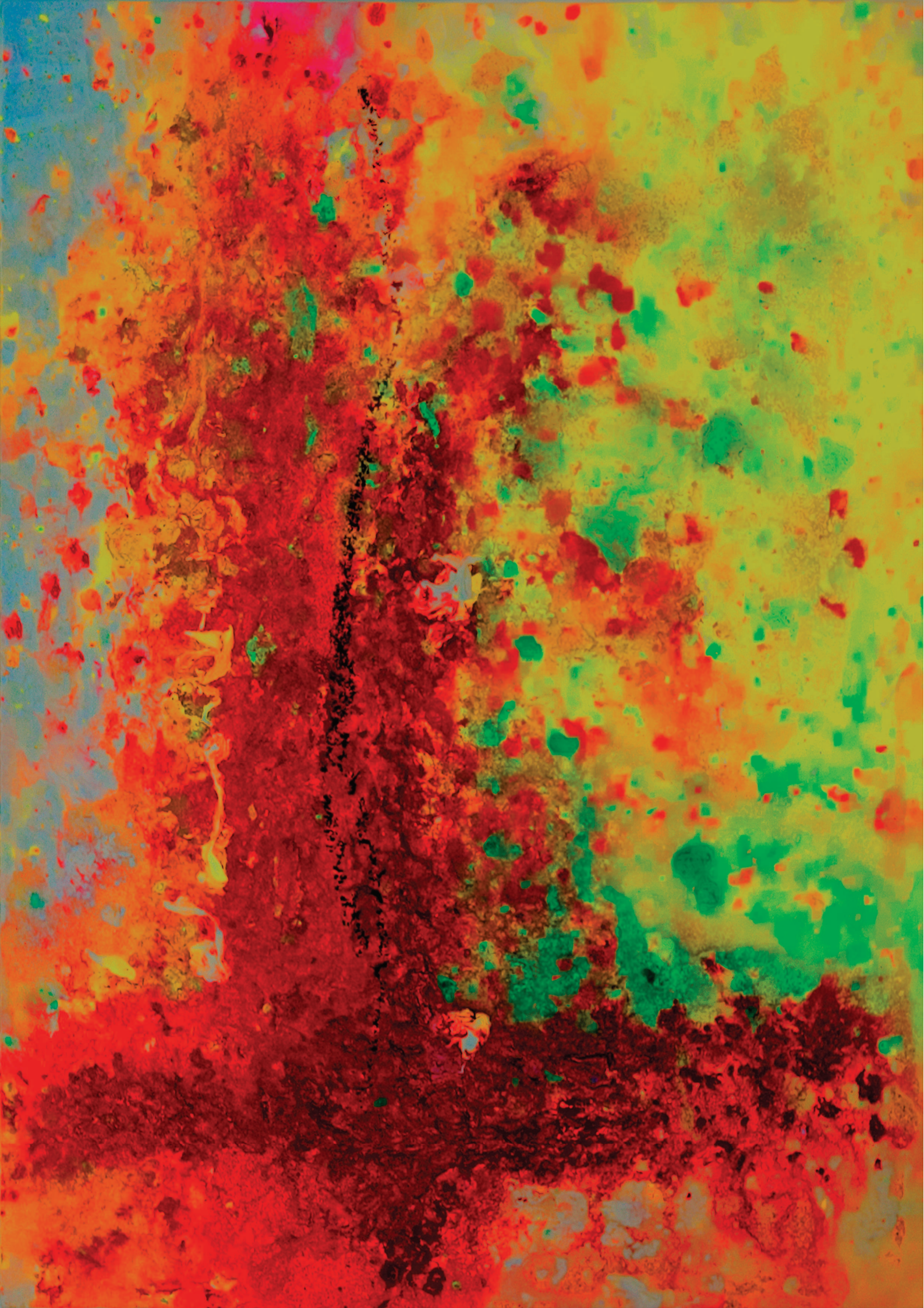
number of features

Relative absolute error on dominance



	114	20	10	5	4	3	2	1
Bayes Net	99.64	99.64	99.64	99.64	99.64	99.64	99.64	99.64
Naive Bayes	97.98	93.32	81.41	74.10	84.10	76.23	79.62	84.02
Multilayer Perceptron	100.41	99.71	100.02	100.05	100.04	100.20	100.27	100.32
Random Forest	100.55	99.35	95.32	95.44	98.58	97.10	86.72	93.76
Naïve Bayes Tree	83.75	96.06	92.19	99.03	96.36	96.36	89.44	94.08
Decision Stump with Ada Boost	103.29	103.43	102.73	102.66	103.11	103.11	101.02	101.84

number of features



	114	20	10	5	4	3	2	1
Bayes Net	99.22	99.22	99.21	99.21	99.21	99.21	99.21	99.21
Naive Bayes	108.98	100.49	80.32	76.60	68.31	68.87	78.42	84.31
Multilayer Perceptron	98.92	100.00	99.69	99.70	99.71	99.61	99.79	99.76
Random Forest	94.92	86.37	82.22	79.43	76.00	67.38	75.21	73.35
Naïve Bayes Tree	106.05	91.45	89.35	85.27	83.48	82.34	79.83	76.27
Decision Stump with Ada Boost	102.45	92.77	87.56	87.52	87.52	87.50	87.49	87.49