RESEARCH　　　　　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

CrossMark

Multi-domain and multi-task prediction of extraversion and leadership from meeting videos

Ahmet Alp Kindiroglu[1*] ⓘ, Lale Akarun[1] and Oya Aran[2]

Abstract

Automatic prediction of personalities from meeting videos is a classical machine learning problem. Psychologists define personality traits as uncorrelated long-term characteristics of human beings. However, human annotations of personality traits introduce cultural and cognitive bias. In this study, we present methods to automatically predict emergent leadership and personality traits in the group meeting videos of the Emergent LEAdership corpus. Prediction of extraversion has attracted the attention of psychologists as it is able to explain a wide range of behaviors, predict performance, and assess risk. Prediction of emergent leadership, on the other hand, is of great importance for the business community. Therefore, we focus on the prediction of extraversion and leadership since these traits are also strongly manifested in a meeting scenario through the extracted features. We use feature analysis and multi-task learning methods in conjunction with the non-verbal features and crowd-sourced annotations from the Video bLOG (VLOG) corpus to perform a multi-domain and multi-task prediction of personality traits. Our results indicate that multi-task learning methods using 10 personality annotations as tasks and with a transfer from two different datasets from different domains improve the overall recognition performance. Preventing negative transfer by using a forward task selection scheme yields the best recognition results with 74.5% accuracy in leadership and 81.3% accuracy in extraversion traits. These results demonstrate the presence of annotation bias as well as the benefit of transferring information from weakly similar domains.

Keywords: Personality predicition, Transfer learning, Multi-task learning, Meeting analysis, Video blogs, Computer vision, Social computing

1 Introduction

The personality of a human is highly correlated with his behavior and decisions in work and leisure environments. Through experience, humans can become quite good at judging the personality of their contacts and modulating their behavior. Automated systems would benefit from the availability of an automated feedback mechanism based on the prediction of the personality of a user of the system.

Training personality predictors based on non-verbal audio-visual cues is a relatively new and scarcely explored territory in research. One major obstacle holding back research in this field is the lack of large amounts of annotated data. Previous work on the subject used different approaches such as finding co-occurrence features [1] or incorporating similar data from different contexts [2]. However, the success of such systems often relies on jointly utilizing information from different sources and features as recognizing personality based on a low number of extracted features often do not yield high accuracy.

By their construction in psychology, the Big-Five personality traits are supposed to be uncorrelated [3]. However, as Vinciarelli notes in [4], this is not often the case as raters' cognitive and cultural biases tend to create correlations among traits while annotating. In addition, personality traits that can be observed in group settings such as emergent leadership and dominance highly correlate with the Big-Five personality traits, in particular with extraversion. To handle such correlation when building predictive models, we explore the usage of a multi-task learning (MTL) framework to search for a combination of

*Correspondence: alpkindiroglu@gmail.com
[1]Department of Computer Engineering, Bogazici University, Istanbul, Turkey
Full list of author information is available at the end of the article

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77

Page 2 of 14

tasks and features that highly correlate with the leadership and extraversion traits of humans.

Multi-task learning (MTL) has recently attracted extensive research interest in the data mining and machine learning community. It has been observed that learning multiple related tasks simultaneously often improves modeling accuracy and leads to better feature selection, especially in cases where each task has very limited number of training samples. Recognizing the fact that not all tasks are uniformly related, there is substantial research interest in modeling task relationships in the state-of-the-art MTL methods.

In this paper, we present a framework for recognizing extraversion and leadership from meeting videos of the ELEA corpus. Prediction of extraversion has attracted the attention of psychologists because it is able to explain a wide range of behaviors, predict performance, and assess risk [5]. Prediction of emergent leadership, on the other hand, is of great importance for the business community [6]. Therefore, we focus on the prediction of extraversion and leadership since these traits are also strongly manifested in a meeting scenario through the extracted features. In previous studies, Sanchez-Cortes et al. [7] demonstrated the correlation of personality traits with the non-verbal features extracted from the ELEA corpus. Kindiroglu et al. [2] showed that a transfer learning approach leads to an increase in extraversion recognition performance, through the incorporation of another database (i.e., the VLOG database [8]) in the learning framework. Building on these previous work, we argue that the two traits, extraversion and leadership, are related traits and a recognition framework for these traits could benefit from a joint learning approach. Our main contributions in this study are summarized as follows:

- We investigate the application of multi-task learning to the problem of personality prediction, which, to the authors' knowledge, is the first of such study in the literature. We investigate in detail whether the prediction of extraversion and leadership can benefit from multi-task learning and under which conditions. We make use of regularized regression-based multi-task learning to simultaneously learn to predict different personality traits to maximize the recognition performance of leadership and extraversion.
- We compare the results with single-task learning algorithms with various feature selection methods to demonstrate the effectiveness of learning several personality traits together.
- Furthermore, we explore and contrast the benefits of transferring more data from other domains together with multi-task learning to improve the prediction models hampered by the lack of large amounts of data.

In Section 2, we present the related work on personality prediction and multi-task learning. In Section 3, we describe the datasets employed and their annotations. In Sections 4 and 5, we describe the multi-domain learning strategies used in this study and discuss the results. Finally, we present our conclusions in Section 6.

2 Related work

In this section, we review related work on personality prediction from visual material and multi-task learning.

2.1 Prediction of personality traits

Personality is defined as the combination of affects, behaviors, cognitions, and desires that characterizes individuals in unique ways [9]. In the literature, many approaches are used to model personalities of humans. The Big Five model proposed by psychology entails five dimensions of personality that have been used extensively in the literature to describe unique individuals [3]. The personality traits in the Big Five scheme are defined as follows:

- **Agreeableness**: friendly/compassionate vs. analytical/detached
- **Conscientiousness**: efficient/organized vs. easy-going/careless
- **Emotional Stability**: sensitive/nervous vs. secure/confident
- **Extraversion**: outgoing/energetic vs. solitary/reserved
- **Openness to experience**: inventive/curious vs. consistent/cautious

In addition to personalities defined for generic purposes, traits that are extensions of personalities such as emergent leadership, dominance, and competence are significant where interaction between people are critical for success in undertaken tasks. Emergent leadership appears mostly in newly formed groups. The behavior of a participant makes a person a good leader or not, without considering past information of competence, related task performance or friendship. This trait highly correlates with dominance and the Big Five personality traits [7]. In addition, Sanchez-Cortes et al. in [10] demonstrate that correlations exist between the perception of leadership and automatically extracted visual cues.

Different methods are used to evaluate personalities of people, the most common being questionnaires. The most widely used are NEO and TIPI questionnaires with 240 and 10 questions [11, 12]. For tasks like automatic recognition, such questionnaire are used to generate ground truth annotations. These annotations are either generated by self-reporting of the user or the reporting of an observer. While self-reporting presents a more accurate analysis of a persons' personality, subjects tend to

Kindiroglu *et al. EURASIP Journal on Image and Video Processing*   (2017) 2017:77

Page 3 of 14

bias their answers towards more socially acceptable if the conducted test results may affect the person in a negative manner (i.e., failing a job interview) [12]. Therefore, obtaining results from multiple observers is a common approach in evaluating personalities.

In the literature, the automatic recognition of personality from various multimedia-based sources is a field of study that is gaining attention. One of the earliest studies on automatic personality prediction was done by Mairesse et al. [13]. The study explored the effectiveness of textual and speech-based acoustic features in recognizing the Big Five traits. More recent studies on the topic focused on recognizing personality from a greater amount of data sources covering human interactions. These are text analysis-based methods [13–15] focusing on choice of words, acoustic-based methods focusing on non-verbal as well as verbal parts of speech [16, 17], wearable devices visual analysis-based methods from image and videos [8, 18–20]. A more complete list of personality prediction literature can also be found in the survey of Vinciarelli et al. [4].

In the automatic personality recognition research literature, the datasets and features pointing at personalities show a large amount of diversity in terms of their data domains. These studies use the abundance of personal data from different domains such as short essays, blog posts [21], social media profiles [18], surveillance with wearable sensors [22, 23], computer game behavior [24] and videos containing interaction between people [10, 17, 20].

One of the major efforts on automatic Big Five personality trait recognition occurred with the Chalearn Looking at People challenge. The challenge organized by Lopez et al. [25, 26] had 10000 15-s job interview videos with continuous Big Five annotation scores in the range of [0, 1]. The challenge received many contributions using deep learning methods for feature extraction and personality prediction. The top group used a combination of CNN- and LGBTOP-based scene and face features with extreme learning machines [27]. Other groups used multi-modal LSTM neural network architectures to better represent the spatiotemporal nature of personality recognition [28]. Some of the audio-visual features, first used in this study, were also utilized in this challenge and achieved scores that were comparatively close to these state-of-the-art results with a simple single-task random forest classifier [29].

2.2 Multi-task learning
Multi-task learning methods in the literature work in a supervised classification or regression framework where models for similar tasks are learned simultaneously. These methods focus on improving the overall prediction performance by sharing information between tasks. Multi-task

learning-based approaches assume that similarities and differences exist among different tasks, and they employ two different approaches to jointly model them. Earlier work on the topic makes use of a lasso-like regularized regression framework. Evgeniou et al. [30] first made use of regularized regression concerned with capturing these shared structures among tasks.

Following on the regularized regression approach, many different models with different regularization methods were proposed. Among these methods, Argyriou et al. [31, 32] tried to find a shared low-dimensional representation for all tasks using L_{21} norm regularization and used these shared features to find task-specific learners for each class. In [33], Gong et al. propose using the L1 norm to find outlier tasks while simultaneously using it on rows to find shared features. They then remove the outlier tasks and perform L_{21} norm multi-task learning on the clean dataset composed of similar tasks. In Jalali et al.'s work [34], the sum of two matrices are used to represent the parameters and these are regularized differently to learn both shared features and individual outliers for different tasks separately.

Follow-up works extended and generalized these concepts by focusing on issues such as learning the relatedness of tasks. Some examples include the work of Kang et al. [35], who used a clustering based approach with the MTL regularization method to group similar tasks together, and Chen et al., who used a graph based structured regularization approach to encode the similarity information between tasks [36].

3   Data and annotations
We perform personality impression prediction on multi-person meeting videos of the publicly available ELEA (Emergent LEAder) corpus. The corpus includes 27 separate group meetings of 3 or 4 members [7]. This dataset includes audio and visual data and personality traits scored by external observers, such as the Big Five personality traits, and by group members, such as the perceived leadership, dominance, competence, likeness, and ranked dominance.

In the videos of this dataset, the participants take part in a winter survival game. As the survivors of an airplane crash, the participants were asked to rank the importance of 12 items they would take with them to increase the chance of survival. After ranking the items individually, the participants ranked them as a group. The participants discuss and try to convince each other while being seated around a table. The entire interaction scene is recorded via a microphone array and wide-angle web cameras.

Personality impressions for the Big Five traits were collected from external observers using the Ten Item Personality Inventory (TIPI), with a 7-point Likert scale [11]. Each annotator watched a 1-min segment of a

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77

Page 4 of 14

participant from the meeting, which corresponds to the segment that includes the participant's longest speaking turn. Three different annotators annotated each video and a total of five annotators annotated the whole dataset. In addition to the Big Five traits, the dataset also includes annotations for the perceived traits as annotated by the group members at the end of the meeting. These traits include perceived leadership, dominance, competence, liking, and ranked dominance. We will call these traits as the leadership traits in the rest of the paper. More details on the dataset and the annotations can be found in [37] and in [38].

In this paper, we used a subset from the ELEA corpus. The subset consists of audio-visual (AV) recordings of 27 meetings. The extracted videos of meetings contain long monologues by a single user. These segments contain parts of the meeting where the user is actively speaking.

In order to infer the personality and leadership traits, we extract different kinds of non-verbal audio-visual features as listed in Table 1.

The second dataset we have used is the publicly available IDIAP VLOG dataset which contains the video blogs (vlogs) of video bloggers (vloggers) downloaded from YouTube. It contains a total of 404 videos. The dataset includes Big Five personality annotations obtained by crowdsourcing. The personality annotations were collected via crowdsourcing on Amazon Mechanical Turk, using the TIPI questionnaire. Five different annotators annotated each vlog, after watching the first conversational minute of each vlog. More details about the dataset and the annotations can be found in [8]. Since it is recorded in a different social interaction setting, where the vlogger is the only participant, it is not possible to extract the perceived leadership traits as in meeting videos.

The extracted features include visual features based on the subject's motion. The dataset contains a large number

Table 1 List of non-verbal features extracted from the ELEA dataset

Feature ID	Features	Description
1:4	Speaking status	Speech segments of the target person
5:8		Interruptions
9:18	Prosodic speech features	Pitch
19:28		Energy
29:32	Visual activity	Head motion
33:37		Body activity
38:59	wMEI	Motion template-based visual activity features
60:73	Visual focus of attention	Visual attention features

of samples for a small subset of features and tasks present in the ELEA corpus. The overlap of common tasks and features for the ELEA and VLOG datasets can be visualized in Fig. 1.

4 Methodology

In our experiments, we aim to perform recognition on the extraversion and leadership traits using the extracted audio-visual features from Table 1. For each subject in the dataset, the ground truth scores for the personality traits were obtained by averaging the personality scores coming from each annotator. For the ELEA dataset, each video is annotated for the Big Five personality traits by three different annotators. The leadership traits on the ELEA dataset were annotated by the fellow group members. It means either two or three annotators, excluding the subject, depending on the group size. For the VLOG dataset, each video is annotated for the Big Five personality traits by five different annotators. These traits were binarized by using the median values for each trait, thus creating a set of labels suitable for binary classification.

In the initial single-task prediction context, a task is defined as the learning of a personality trait. As there are two personality traits, we have two tasks. We consider each personality trait as a single task and classify them separately. For the baseline method, extraversion and leadership classification models were constructed using support vector machine, decision forest and ridge regression classifiers. A leave-one-out cross-validation scheme was used to evaluate prediction accuracy of each task.

In order to improve the reliability of the extraversion and leadership trait predictions, we focus on incorporating data by learning several tasks together via multi-task learning. Previous studies on the ELEA dataset demonstrate that incorporating knowledge from other domains were beneficial for predicting the extraversion trait [39]. However, the non-availability of leadership-based annotations in the VLOG domain reduces the effect of the transfer for the leadership prediction task. In order to overcome such a problem, we utilize a multi-domain and multi-task learning framework where we incorporate the five personality and four leadership traits to improve recognition performance in the extraversion and leadership recognition tasks. The flow of the baseline, multi-task, and multi-domain and multi-task approaches used in this study can be observed in Fig. 2.

In the next sections, we give the background and the details of the methods that we use for feature selection transfer learning and multi-task learning.

4.1 Feature selection

The objective of feature selection is to eliminate noisy, irrelevant and redundant features to obtain shorter feature vectors, where better classification is possible. This allows

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77
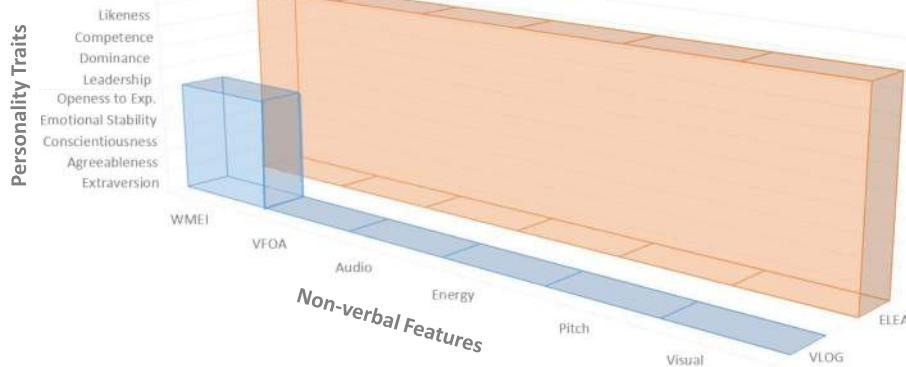
Page 5 of 14

Fig. 1 Set of personality traits and features used in prediction for the emergent leader analysis (orange) and VLOG (blue) tasks. VLOG segment contains a smaller number of features and tasks

the building of better performing models with regard to both accuracy and speed while also presenting a better understanding of the data and the prediction process.

In the feature selection literature, there are two popular methods: filtering-based and wrapper-based methods [40]. Filtering-based methods, which are more cost efficient, rely on extracting and comparing the inherent properties of features such as relevance, correlation, and mutual information. Wrapper-based methods, on the other hand, make use of black-box learning algorithms to predict learner performance using different search strategies.

In filtering-based approaches, finding the most efficient subset of labels given a set of data and its class labels

is possible through a method called maximum dependency. Using mutual information as a measure of similarity such an approach can calculate the best possible subset of features which have the largest dependency on the label of the given class. However such an approach requires the calculation of multivariate density functions for $p(x_1, \ldots, x_m)$ and $p(x_1, \ldots, x_m, c)$ which are difficult in high dimensional space.

In this study, we use maximum relevance minimum redundancy (MRMR), which is a frequently used feature selection algorithm. It is equivalent to the max-dependency algorithm for first-order search. In MRMR, the mean value of all mutual information values across individual features and class labels are calculated as the
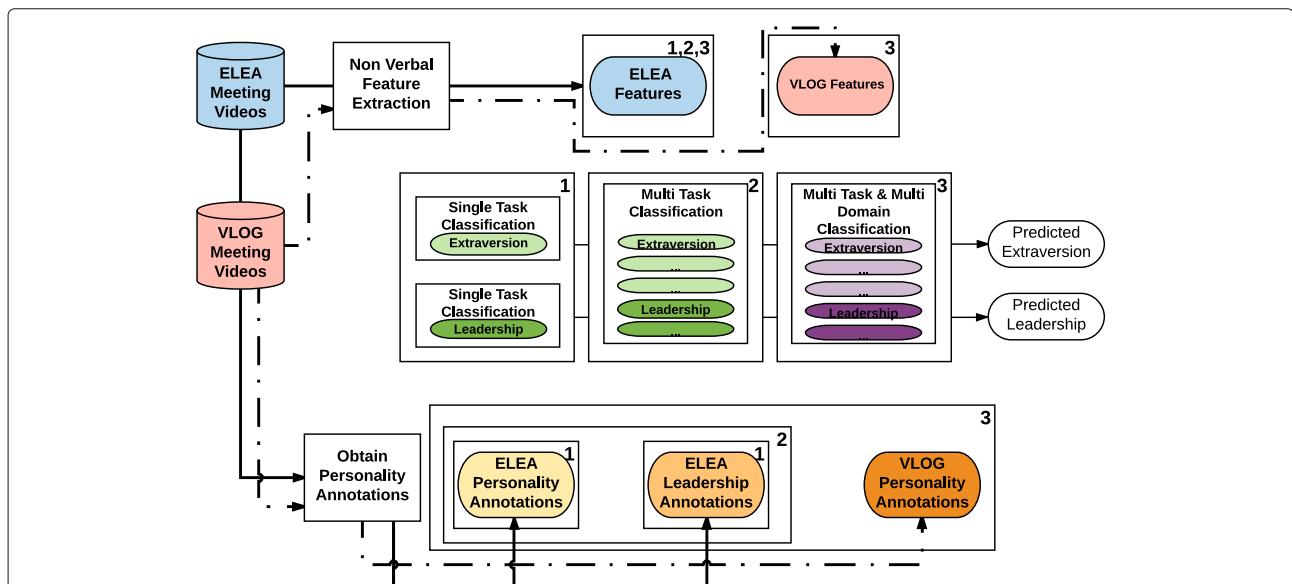

Fig. 2 The features and annotations obtained from the ELEA and VLOG datasets and the approaches that use them are depicted by the method numbers on the encapsulating boxes. The baseline method is designated as method 1, the multi-task method is method 2, and the multi-domain and multi-task approach is designated as method 3

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77

Page 6 of 14

max-relevance values. The minimal redundancy between variables is calculated by finding the mutual information scores between variables. Finally, by using a function such as the difference of max relevance and min redundancy, the most dependent feature can be selected. Using this method in an incremental search, it is possible to find the most important features.

4.2 Multi-task learning

Multi-task learning is a statistical spatial feature mapping framework where identical input data from different tasks are jointly learned. Joint learning of unrelated tasks can often lead to more informative sparse descriptors. Such descriptors can provide better discriminative power for the unusual features of the data. When the tasks are uncorrelated, their contribution can act as noise to other tasks thus improving generalization. Adding tasks may also increase the weight of parameters with more tasks allowing a better capture of the feature space.

In our formulation, we have a set of non-verbal features extracted from the ELEA corpus videos, represented as $X = x_{1..N}$, where each i represents the features from a single video sample. These are the same among tasks $t = 1..T$. The labels for each task are denoted as $Y(t)$. The aim of our framework is to learn a dictionary W which maps the training samples, X to their labels, Y. Classical machine learning frameworks aim to learn a joint dictionary defined as $Y = WX$. Solutions to such problems can be obtained by using least-square-based methods. Desired properties of the dictionary can be obtained by adding a regularization term such as the element-wise $L1$ matrix norm as imposed by the LASSO algorithm [41].

$$\operatorname*{argmin}_{W} ||WX - Y||_F + \alpha ||W||_{1,1} \qquad (1)$$

$$where ||W||_{p,q} = \left[\sum_{j=1}^{n} \left(\sum_{i=1}^{m} |a_{ij}|^p \right)^{q/p} \right]^{1/q}$$

As we can see from Eq. 1, the standard LASSO does not distinguish the inputs and regression coefficients from different tasks. Multi-task learning algorithms aim to minimize the sum of squared errors (SSE) for each task rather than finding an optimized SSE for all training samples, X.

4.2.1 Multi-task LASSO

The multi-task LASSO(MTL) method [42] is an extension of the regular LASSO method proposed by Tibshirani et al. [41]. The multi-task LASSO allows fitting multiple regression problems jointly enforcing the selected features to be the same across tasks.

$$\operatorname*{argmin}_{W} \sum_{i=1}^{t} \left|\left| W_i^T X - Y \right|\right|_F + \alpha_1 ||W||_{1,1} + \alpha_2 ||W||_F \quad (2)$$

As formulated in Eq. 2, regularization parameters of MTL are α_1 and α_2, which control the sparsity of the dictionary matrix and the norm penalty respectively.

4.2.2 $L_{2,1}$ norm regularized multi-task LASSO

While MTL enforces a sparse representation on the entire dictionary, there is no emphasis on extracting the common features among the tasks. To enforce extracting shared features among tasks, Argyriou et al. [31], suggests adding a $L_{2,1}$ regularization term to the regularization problem. Since L_1 norm favors sparsity and the L_2 norm favors uniformity, the $L_{2,1}$ provides the desired parameter sparsity shared across tasks by summing the Euclidean norm of each column.

$$\operatorname*{argmin}_{W} \sum_{i=1}^{t} \left|\left| W_i^T X - Y \right|\right|_F + \alpha_1 ||W||_{2,1} + \alpha_2 ||W||_F \quad (3)$$

4.2.3 Trace norm regularized multi-task LASSO

Another multi-task learning method making use of regularized regression is the trace norm regularized method proposed by Chen et al. [43]. The goal in trace norm regularized lasso is to find a low-dimensional subspace shared by different tasks. In the regularization, trace norm is used as a rank function.

The trace norm of the matrix $||W||_*$ is calculated by summing its eigenvalues. Minimizing the regression objective function in Eq. 4 yields W matrices of minimal rank for each task. The parameter α_1 is used to control the rank of W.

$$\operatorname*{argmin}_{W} \sum_{i=1}^{t} \left|\left| W_i^T X - Y \right|\right|_F + \alpha_1 ||W||_* \qquad (4)$$

4.2.4 Robust multi-task learning

Robust multi-task learning (ROBUST) algorithm focuses on the elimination of outlier tasks in the Multi-Task Lasso Framework. The model assumes that the model matrix W can be decomposed into two matrices that capture task relatedness and group sparsity.

$$\operatorname*{argmin}_{W} \sum_{i=1}^{t} \left|\left| W_i^T X_i - Y_i \right|\right|_F + \alpha_1 ||L||_* + \alpha_2 ||S||_{1,2} \quad (5)$$

$$subject\ to : W = L + S$$

In Eq. 5, the trace norm enforces a low-rank structure to couple tasks that are closely related, while the $L_{1,2}$ norm which is the row grouped L_1 norm captures outlier tasks.

4.2.5 Robust multi-task feature learning

Robust multi-task feature learning (RMTFL) algorithm focuses on the elimination of outlier tasks in the $L_{2,1}$ norm regularized multi-task LASSO framework [33]. In this framework, the W matrix is decomposed into two

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77

Page 7 of 14

components as P and Q. The formulation of the method is provided in Eq. 6.

$$\underset{W}{\operatorname{argmin}} \sum_{i=1}^{t} \left\| W_i^T X - Y \right\|_F + \alpha_1 \|P\|_{2,1} + \alpha_2 \left\| Q^T \right\|_{2,1} \quad (6)$$
$$subject\ to: W = P + Q$$

In this equation, the P matrix captures group sparsity by jointly selecting sparse features across tasks while the Q matrix captures jointly selected sparse samples thus capturing outlier tasks. The parameters α_1 and α_2 control the effect of each parameter in the learning process respectively.

4.2.6 Dirty multi-task LASSO
While the MTL and L_{21} norm regularized MTL performs well with ideal data, the data may not always be represented ideally using a single structure. Jalali et al, in [34], propose to decompose the model W into two components as P and Q, where one captures shared features among tasks while the other captures intrinsic properties that are useful in recognizing individual tasks.

$$\underset{W}{\operatorname{argmin}} \sum_{i=1}^{t} \left\| W_i^T X - Y \right\|_F + \alpha_1 \|P\|_{1,\infty} + \alpha_2 \|Q\|_{1,1}$$
$$subject\ to: W = P + Q$$

As formulated in Eq. 7, P captures group sparsity through $L_{1,\infty}$ norm which sums the maximum value of each row. Q enforces sparsity on the overall structure of the data and both P and Q are subject to $W = P + Q$.

4.2.7 Multi-task discriminant analysis
Multi-task discriminant analysis (MTDA) [44] can be seen as a multi-task extension of the widely used supervised dimensionality reduction technique linear discriminant analysis (LDA) (Fukunaga 1991). Instead of simply pooling the data for multiple learning tasks together and learning a common transformation for all tasks, MTDA learns a separate transformation for each task.

Each MTDA transformation consists of two parts, one specific to the corresponding task and one common to all tasks. MTDA is based on an objective function, seen in Eq. 7, which is similar to that of the single-task LDA.

$$\underset{W_i}{\operatorname{argmax}} \frac{tr \left(W_i^T S_b^i W_i M \right)}{tr \left(W_i^T S_t^i W_i M \right)} s.t. W_i^T W_i = I_{d'} \quad (7)$$

While most existing multi-task learning methods can only handle learning tasks with data sharing the same feature space, MTDA can deal with heterogeneous feature spaces, allowing the incorporation of data with missing features such as the VLOG dataset.

4.3 Transfer learning
Many machine learning methods work well under certain assumptions such as having enough labeled data for each aspect of the features they model. In reality, such assumptions are usually not satisfied leading to imperfect models as is the case with the ELEA dataset [38]. The goal of transfer learning is to improve learning in a target task by utilizing knowledge from other source tasks.

In our previous study [2], the effect of transfer in personality affect prediction was analyzed in ELEA and VLOG datasets. Baseline machine learning approaches were used in a transfer framework with target-only, source-only, combined, and multi-task LDA-based learning methods.

In this study, we employed different baseline- and transfer-based approaches for domain adaptation. Since the task of personality prediction used a shared set of features in both datasets, using multi-task learning methods to jointly learn from both datasets was possible. We made use of the regularized regression based multi-task learning approaches using five personality traits and four leadership traits from the ELEA dataset and five personality traits from the VLOG dataset as tasks.

5 Results and discussion
We predict two personality traits, extraversion and leadership from meeting videos. The lack of sufficiently large amounts of annotated data necessitates the use of transfer and multi-task learning approaches to augment the feature space and to preprocess the data. Experimental results are presented for multi-task learning and transfer learning frameworks. Various multi-task and transfer learning methods are used to augment the feature space and preprocess the input data. The need for both approaches arises from the lack of sufficiently large amounts of annotated data.

In our framework, multi-task learning approaches are employed to allow learning of multiple personality traits together in order to capture sparser and better representations on personality prediction models. On the other hand, transfer learning approaches allow us to use features extracted from other domains.

In the baseline recognition system, the learning task is accomplished by support vector machine [45], decision forest, and ridge regression algorithms. LibSVM implementation is used for SVM and the C parameter of the SVM is selected from the $\left[2^{-5}, 2^5 \right]$ range. Matlab's treebagger algorithm is used for performing learning with decision forests with 100 trees. Parameters for both methods were optimized using a two-layered cross-validation scheme. In the ridge regression classifier, we used the 0/1 labels as the scores and estimated the label based on thresholding the estimated score at 0.5. Both source and target data were z-normalized, separately using each datasets' training

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77

Page 8 of 14

samples. The ridge parameter values were selected from a range of [2, 150]

5.1 Transfer learning framework

In the transfer learning literature, domains are labeled as source and target domains according to their role in the learning process. Since we are interested in increasing the prediction accuracy of extraversion in small group settings, we choose the ELEA dataset as our target and the VLOG dataset as our source domains. We used the VLOG corpus to enhance a recognition model trained to recognize videos from the ELEA dataset.

Previous studies with the ELEA corpus using classical approaches obtained limited success due to a lack of sufficient training samples. As a method of easier bootstrapping, additional samples belonging to a similar domain can be utilized. In [46], personality annotations in a video blog dataset were used to transfer information to the training models obtained from the ELEA dataset. However, while beneficial in allowing the creation of better models for extraversion, as a result of the differences between video blog and group meeting domains, the transfer was only possible on the Big Five personalities and with visual WMEI features.

In our experiments, a leave-one-out cross-validation scheme was used to explore the effect of target ELEA training data on transferred recognition performance. The target data, chosen as the leadership and extraversion traits of the ELEA dataset, is divided into 102 folds for 102 samples. Models using the five Big Five personality and five leadership annotations from the ELEA dataset and five Big Five personality annotations of the VLOG dataset are used as external data sources. Learning models are enhanced using the multi-task learning algorithms from Section 5.2. We report the final accuracy on the ELEA dataset.

5.2 Multi-task learning framework

We used the samples of the ELEA dataset in a leave-one-out cross-validation framework to obtain multi-task recognition results. In our experiment sets $D = \{X, Y\}$ we have 10 sets of binary labels, Y for different tasks and only one set of features X extracted from 1-min video segments. The implementations of the methods are based on the MALSAR package [42].

In order to find the best performing multi-task method alongside the best performing learning algorithm, we attempt to find the best performing feature/multi-task algorithm/learner combination with their different parameters.

In our experiments, we use seven different multi-task approaches as summarized in Table 2.

These approaches are all utilized as feature extraction algorithms to obtain a new feature space representation.

Table 2 List of classifiers used in the study

	Name	Acronym
Regularized regression classifiers	Multi-task lasso	LASSO
	$L_{2,1}$ norm regularized with least squares loss	L21
	Trace norm regularized regression	TRACE
	Robust multi-task learning	ROBUST
	Robust multi-task feature learning	RMTFL
	Dirty multi-task learning	DIRTY
Multi-task feature learning	Multi-task discriminant analysis	MTDA

During training, we use a validation set to search for optimal parameters. In addition to searching for the best combination of learners and feature extraction methods, we also search for the multi-task and learning algorithm parameters.

5.3 Experiments

In our experiments, we test the prediction accuracies of different features, different classifiers, multi-task learning methods, and transfer learning methodologies. We perform extensive tests and try to answer several research questions. In order to interpret the results better, we have organized our tests around the following questions:

5.3.1 How difficult is the prediction of individual personality traits?

In the ELEA dataset, 10 personality traits have been annotated. In a meeting scenario, some traits are more strongly manifested than others. On the other hand, we make use of extracted audio and visual features to predict perceived personality traits. What will be the success of this prediction? To test this, we use all available features and use three state-of-the-art classifiers: support vector machine (SVM), ridge regression (RIDGE), and random forest (RF). In Table 3, our tests illustrate that the prediction accuracy of extraversion and leadership are relatively high. Using the ridge regression method, leave-one-out cross-validation scheme demonstrates recognition accuracies of 75.5% for the extraversion trait and 68.6% for the leadership trait. It is observed that the recognition performances of the other personality traits are lower. Therefore, we regard them as complementary tasks rather than focusing on their recognition. We predict only extraversion and leadership in further tests and use the other tasks as complementary information in multi-task learning.

5.3.2 How useful are the individual features?

Next, we make use of the individual groups of features extracted from the ELEA dataset to analyze overall recognition performance of groups of features. We make use of

Kindiroglu *et al. EURASIP Journal on Image and Video Processing*   (2017) 2017:77

Page 9 of 14

Table 3 Single-task prediction accuracies for personality traits

	Personality traits									
	Ldr	Dom	Cmp	Lkn	RDom	Ext	Agr	Con	Emo	Opn
SVM	64.7	68.6	57.8	55.9	60.8	74.5	58.8	56.9	60.8	53.9
RIDGE	68.6	53.9	54.9	64.7	55.9	75.5	55.9	60.8	58.8	52.0
RF	68.6	59.8	55.9	55.9	54.9	73.5	52.0	51.0	50.0	52.9

three baseline two-class classifiers (SVM, RIDGE, RF) to evaluate the performance of each feature group as seen in Table 4.

In Table 4, bold figures indicate the best performance in each category. We observe that different features have varying predictive power for extraversion and leadership. Visual features perform well for the prediction of both traits. On the other hand, WMEI features are more important for the prediction of extraversion while VFOA features are more important for the prediction of leadership.

In order to concentrate on individual features rather than a group of features, and test the effectiveness of their combinations, we employ feature selection (MRMR) and extraction (PCA) for a single-task prediction. The results are presented in Table 5. Baseline results with all features are given in the first row for comparison. We observe that feature extraction (PCA) is not beneficial as it decreases the recognition accuracy. Feature selection, on the other hand, increases performance: It is observed that the peak performances for extraversion prediction are obtained at a subset size of 40 and 50 features using MRMR and SVMs and 50 features with RDFs. For leadership prediction, the peak results are given using MRMR with random decision forests using 30 and 40 features while Ridge regression achieves a good performance of 71.6% using only 10 features. Exploration of the selected features for each trait demonstrates that while features for a leadership prediction trait are mostly selected from speaking status and visual focus of attention features, features for extraversion

prediction are selected from nearly all the available feature sources.

5.3.3   How beneficial is multi-task learning?

As described in Section 2, there are correlations in perceptions of personality traits and this leads to correlations in the annotated data. This leads us to hypothesize that we can increase the prediction accuracy of extraversion and leadership using other personality traits. To test this hypothesis, we have conducted experiments on the ELEA corpus as summarized in Table 6. We use five personality and five leadership traits as different tasks to jointly learn prediction models and use them with nine different classifiers. Since only leadership and extraversion traits correlate highly with the features, we extracted in the ELEA dataset, we observe and report prediction performances only on those two traits. We further include experiments with MRMR and MTDA feature selection approaches to search for the best performing learner combination. We observe that the best accuracies are 78.4% for extraversion and 72.6% for leadership, displaying a 2% increase from the baseline results. The results demonstrate that in comparison to the baseline methods, the regularized regression-based methods do not benefit from the utilized feature selection and preprocessing approaches. This can be explained away by the fact that contrary to single-task methods, these classifiers have the ability to utilize features that are only useful for some tasks and are detrimental to others. The best results are obtained using the

Table 4 Prediction performance for individual feature groups

	Extraversion			Leadership		
	SVM	RF	RIDGE	SVM	RF	RIDGE
Audio	46.1	59.8	54.9	55.9	59.8	55.9
Energy	59.8	62.7	66.7	53.9	57.8	62.7
Pitch	61.8	60.8	59.8	58.8	57.8	57.8
Visual	64.7	66.7	70.6	54.9	54.9	62.7
WMEI	60.8	65.7	62.7	56.9	55.9	61.8
VFOA	52.0	44.1	60.8	64.7	58.8	65.7
All	*74.5*	73.5	*75.5*	64.7	*68.6*	*68.6*

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77

Page 10 of 14

Table 5 Effects of feature selection and extraction approaches for single-task prediction for extraversion and leadership

	# Ftrs	Extraversion			Leadership		
		SVM	RIDGE	RF	SVM	RIDGE	RF
None	73	74.5	*75.5*	72.5	64.7	68.6	67.6
MRMR	10	72.5	71.6	75.5	64.7	*716*	696
	20	72.5	67.6	75.5	57.8	64.7	71.6
	30	72.5	66.7	74.5	57.8	66.7	*72.5*
	40	*76.5*	72.5	75.5	*67.6*	63.7	*72.5*
	50	*76.5*	70.6	*76.5*	*67.6*	58.8	65.7
	60	75.5	73.5	74.5	66.7	65.7	71.6
	70	74.5	73.5	74.5	65.7	69.6	69.6
PCA	10	57.8	66.7	58.8	52.0	64.7	61.8
	20	67.6	72.5	64.7	61.8	63.7	57.8
	30	69.6	67.6	65.7	62.7	58.8	57.8
	40	70.6	74.5	68.6	60.8	60.8	63.7
	50	74.5	70.6	66.7	62.7	66.7	55.9
	60	73.5	69.6	62.7	65.7	64.7	56.9
	70	73.5	62.7	66.7	65.7	57.8	55.9

multitask Lasso method, which solely focuses on sparse features and does not take into account whether they are shared across tasks.

5.3.4 Can we transfer knowledge from a different, richer domain?

We next evaluate the performance of transfer learning. For transfer learning, we use VLOG as the source domain and ELEA as the target domain. The VLOG dataset contains only a subset of the features in ELEA. The common features are the WMEI features. Therefore, only traits that are well represented by WMEI are positively affected by the transfer. Therefore, we focus only on extraversion prediction. The effect of using different methods for

transfer learning for extraversion prediction is illustrated in Fig. 3. We observe that lasso and dirty regularization-based regression methods yield significant improvements in performance, both yielding 78.4% prediction accuracy.

5.3.5 Is knowing too much detrimental?

In multi-domain learning, negative transfer across tasks is a problem. Trying to learn uncorrelated tasks together may often cause a decrease in performance. Therefore, better performances may be obtained by transferring only from relevant domains. We explore combining the best performing tasks in regularized regression-based learning algorithms. For this purpose, we make use of forward task selection and incorporate the most relevant tasks one

Table 6 Multi-task recognition for leadership and extraversion traits by utilizing all personality and leadership traits on ELEA

	Preprocessing:	None		MRMR-45		MTDA	
		Lead	Ext	Lead	Ext	Lead	Ext
Baseline	RF	67.6	74.5	*68.6*	*75.5*	61.8	72.5
	SVM	64.7	74.5	67.6	*75.5*	62.7	*75.5*
	RIDGE	68.6	75.5	58.8	70.6	65.6	70.6
Regularized regression	LASSO	*70.6*	*76.5*	*68.6*	74.5	61.8	74.5
	L21	66.7	72.5	59.8	72.5	64.7	73.5
	TRACE	69.6	72.5	59.8	72.5	*65.7*	71.6
	ROBUST	68.3	72.5	58.8	69.6	56.9	69.6
	RMTFL	69.6	72.5	59.8	69.6	64.7	72.5
	DIRTY	66.3	76.5	59.8	73.5	61.8	73.5

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77
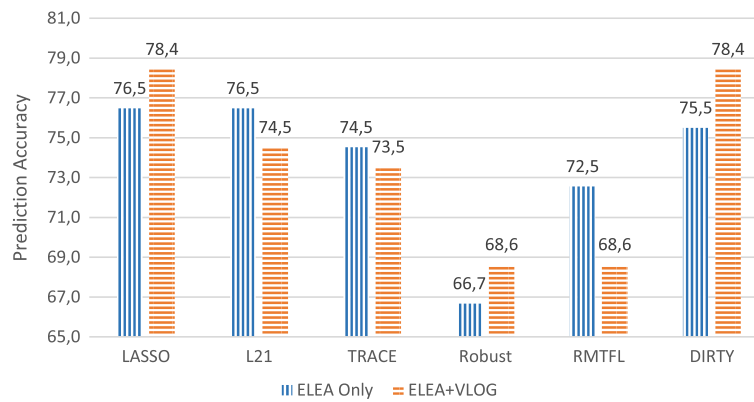
Page 11 of 14

Fig. 3 Prediction accuracy of transfer learning in extraversion prediction

by one. Of the 15 tasks evaluated with forward selection, five are Big Five personality traits, five are emergent leadership traits from the ELEA dataset, and the remaining five are Big Five personality traits from the VLOG dataset.

As seen in Figs. 4 and 5, incorporation of the most correlated tasks at each step of forward selection yields significant increases in the prediction of extraversion and leadership. This increase in performance is more evident in certain methods: L21 shows the best overall performance in both extraversion and leadership prediction. This leads us to search for the tasks that would yield the highest recognition performances when learned together. We observe that among the leadership prediction methods, dirty MTL reaches top leadership prediction performance using two tasks. However, as the number of tasks increases, its performance drops. On the other hand, the recognition performance of L21 and methods that can deal with outliers better such as RMTFL reach

top recognition performance. For extraversion prediction, the best results are given by L21 norm regularized multi task lasso method. However, L21 method performs worse than multi-task lasso when the task count is low and worse than DIRTY MTL when all the tasks are included.

For the trait of leadership, the most beneficial and therefore first incorporated traits in the forward selection scheme were composure, agreeableness, emotional stability from the ELEA, and openness to experiences from the VLOG datasets in the given order. On the other hand, for the extraversion trait, extraversion and conscientiousness from the VLOG domain were selected first, followed by the dominance and likeness traits of ELEA. These results hint that while the largest increases in recognition performance for extraversion come from the VLOG domain data, recognition performance for leadership only benefits from utilizing other tasks of the ELEA corpus.
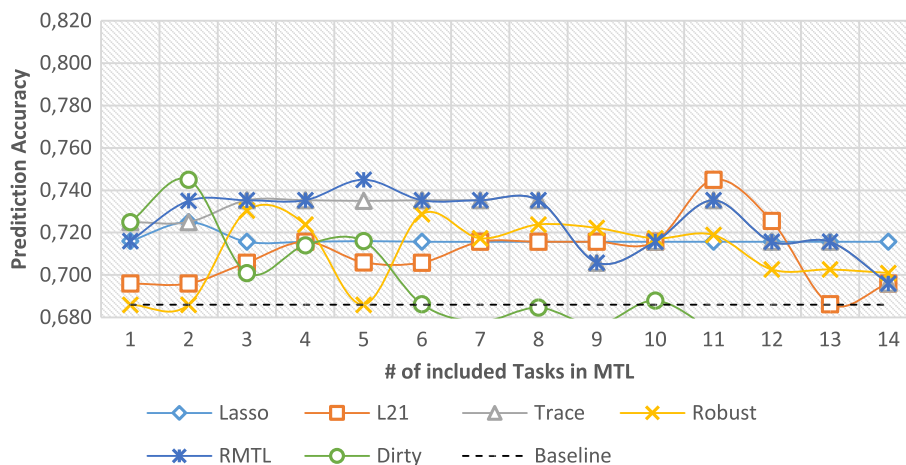

Fig. 4 Leadership. Multi-task learning using forward selection of tasks. At each step, the best performing personality trait is included in the training set. Table presents prediction performance on ELEA dataset for the leadership trait
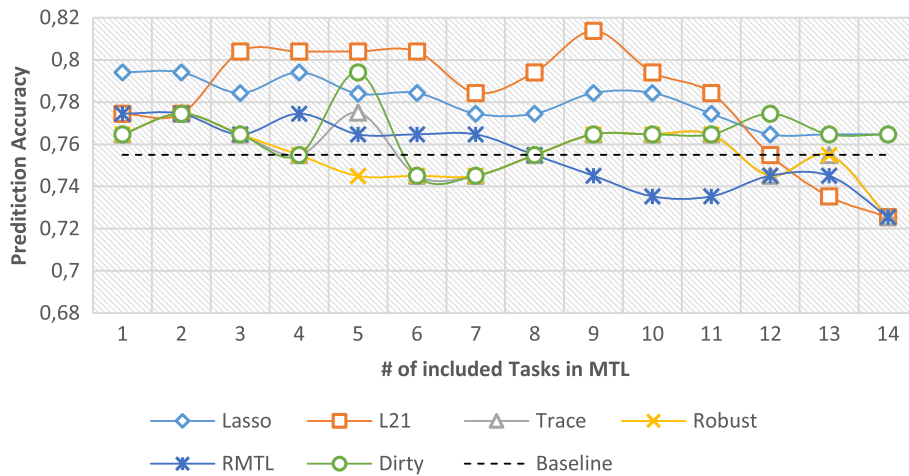
Fig. 5 Extraversion. Multi-task learning using forward selection of tasks. At each step, the best performing personality trait is included in the training set. Table presents prediction performance on ELEA dataset for the extraversion trait

5.3.6 What is the best overall strategy?

Table 7 summarizes the previous work on extraversion and leadership prediction on the ELEA dataset and compares them with the results in this paper. In extraversion prediction, the best results were obtained as 72.5% using WMEI features with transfer learning [2] and as 74.5% using the full set of features used in this study with baseline classifiers [38]. Our baseline classifiers exceeded this by 2% using feature selection approaches. We have obtained almost 6% improvement by using L21 regularized MTL with task selection. Adding transfer learning by itself also leads to about 3% improvement over the baseline. In leadership prediction, the highest results on the ELEA dataset using a similar baseline methodology was reported as 74.5% in [1] using co-occurrence features. Since our features do not include co-occurrence, this result is not directly comparable with ours. In this trait, our baseline classifiers achieved a top recognition performance of 68.6%. Using multitask learning with task

selection, we were able to obtain 74.5% performance, reaching the top performance for the leadership trait on the ELEA dataset that was reported with co-occurrence features.

One interesting result that can be observed is the performance drop of leadership prediction for the MTL method with multi-task LASSO (70.6%) compared to the single-task ridge regression classifier baseline with MRMR feature selection (72.5%). From this, it can be deduced that when using all the tasks, learning with a subset of features is better than using all. However, when the most beneficial tasks are chosen, building regression based classifiers using all the features increases peak performance.

6 Conclusions

In this study, we present a multi-modal prediction framework for extraversion and leadership traits from meeting videos. We explored the usage of several regularized regression based MTL approaches. Using MRMR for feature selection and PCA for feature extraction alongside classifiers such as SVMs, ridge regression, and decision forest, we evaluated our transfer and multi-task learning-based personality prediction framework.

Without the use of transfer learning, the best recognition performance results were obtained by the dirty and robust multi-task lasso methods for leadership and $L_{2,1}$ norm regularized MTL for the extraversion traits. Experiments with transferring data from the VLOG dataset demonstrated an overall increase of 3% for the prediction of both the extraversion and leadership traits. Further use of a forward selection scheme for transferring a choice subset of tasks instead of transferring all tasks further increased the recognition accuracy to 74.5% accuracy in leadership and 81.3% accuracy in the extraversion personality traits. This approach exceeded

Table 7 Performance on the ELEA dataset

	ELEA dataset performance	
	Extraversion	Leadership
Kindiroglu et al. [2]	72.5	–
Aran and Gatica-Perez [38]	74.5	–
Okada et al. [1]	69.6	*74.5*
Baseline—ridge	75.5	68.6
Feature selection and RF	76.5	72.5
MTL—LASSO	76.5	70.6
Transfer—lasso and dirty	78.4	–
Forward selection—L21	*81.3*	*74.5*

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77

Page 13 of 14

the performances reported in [1] by $\sim$ 13% and in [2] by $\sim$ 9%.

The effectiveness of the $L_{2,1}$ norm-based regularization suggests that detecting commonly used features across other personality annotations is a critical aspect in multi-task personality detection in meeting videos. Overall, this indicates that while applying transfer learning and learning multiple traits separately improve automatic leadership and extraversion recognition accuracy, their combination yields a greater increase in creating a better recognition framework.

Availability of data and materials
The datasets supporting the conclusions of this article are available for download from IDIAP dataset access pages at "https://www.idiap.ch/dataset/youtube-personality" and "https://www.idiap.ch/dataset/elea/download-proc" for non-commercial and scientific purposes.

Authors' contributions
OA proposed the main idea, collected the experiment videos, and extracted the features. AK implemented the learning methods and performed the experiments present in this manuscript under the supervision of LA and OA. The authors structured the organization of the manuscript jointly and took part in writing the paper. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Department of Computer Engineering, Bogazici University, Istanbul, Turkey.
[2]IDIAP, Martigny, Switzerland.

References
1. S Okada, O Aran, D Gatica-Perez, in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction - ICMI '15*. Personality trait classification via co-occurrent multiparty multimodal event discovery (ACM Press, New York, 2015), pp. 15–22. doi:10.1145/2818346.2820757. http://dl.acm.org/citation.cfm?id=2818346.2820757. Accessed Sept 2016–Mar 2017
2. AA Kindiroglu, L Akarun, O Aran, in *22nd Signal Processing and Communications Applications Conference (SIU)*. Vision based personality analysis using transfer learning methods (IEEE, Trabzon, 2014), pp. 2050–2053. doi:10.1109/SIU.2014.6830663
3. G Matthews, IJ Deary, MC Whiteman, *Personality traits*, (2003). https://books.google.com.tr/books/about/Personality_Traits.html?id=BqFH_XIq0yYC&pgis=1. Accessed Sept 2016–Mar 2017
4. A Vinciarelli, G Mohammadi, A survey of personality computing. IEEE Trans. Affect. Comput. **5**(3), 273–291 (2014). doi:10.1109/TAFFC.2014.2330816
5. B Lepri, R Subramanian, K Kalimeri, J Staiano, F Pianesi, N Sebe, Connecting meeting behavior with extraversion—a systematic study. IEEE Trans. Affect. Comput. **3**(4), 443–455 (2012)
6. J Kickul, G Neuman, Emergent leadership behaviors: the function of personality and cognitive ability in determining teamwork performance and ksas. J. Bus. Psychol. **15**(1), 27–51 (2000)
7. D Sanchez-Cortes, O Aran, D Gatica-Perez, in *Multimodal Corpora for Machine Learning: Taking Stock and Road mapping the Future*. An audio visual corpus for emergent leader analysis (ACM, Alicante, 2011)
8. J Biel, D Gatica-Perez, The youtube lens: crowdsourced personality impressions and audiovisual analysis of vlogs. IEEE Trans. Multimed. **1520-9210**, 41–55 (2013). doi:TMM.2012.2225032
9. F Alam, E Stepanov, G Riccardi, Personality traits recognition on social network—Facebook (2013). Technical report
10. D Sanchez-Cortes, O Aran, MS Mast, D Gatica-Perez, in *International conference on multimodal interfaces and the workshop on machine learning for multimodal interaction. ICMI*. Identifying emergent leadership in small groups using nonverbal communicative cues (ACM, China, 2010)
11. S Gosling, P Rentfrow, W Swann, A very brief measure of the Big-Five personality domains. J. Res. Personal. **37**(6), 504–528 (2003)
12. B Rammstedt, OP John, Measuring personality in one minute or less: a 10-item short version of the Big Five Inventory in English and German. J. Res. Personal. **41**(1), 203–212 (2007). doi:10.1016/j.jrp.2006.02.001
13. F Mairesse, M Walker, M Mehl, R Moore, Using linguistic cues for the automatic recognition of personality in conversation and text. J. Artif. Intell. Res. **30**(1), 457–500 (2007)
14. L Qiu, H Lin, J Ramsay, F Yang, You are what you tweet: personality expression and perception on Twitter. J. Res. Personal. **46**(6), 710–718 (2012). doi:10.1016/j.jrp.2012.08.008
15. S Argamon, S Dhawle, M Koppel, JW Pennebaker, in *Proceedings of the joint annual meeting of the interface and the classification society of North America*. Lexical predictors of personality type, (Citeseer, St. Louis, 2005)
16. G Mohammadi, A Vinciarelli, Automatic personality perception: prediction of trait attribution based on prosodic features. IEEE Trans. Affect. Comput. **3**(3), 273–284 (2012). doi:10.1109/T-AFFC.2012.5
17. F Valente, S Kim, P Motlicek, in *Proceedings of Interspeech 2012*. Annotation and recognition of personality traits in spoken conversations from the AMI Meetings Corpus, (2012). http://infoscience.epfl.ch/record/192686. Accessed Sept 2016–Mar 2017
18. F Steele, D Evans, RK Green, in *Proc. AAAI Int. Conf. Weblogs Social Media*. Is your profile picture worth 1000 words? Photo characteristics associated with personality impression agreement, (AAAI, California, 2009)
19. R Srivastava, J Feng, S Roy, S Yan, T Sim, in *Proceedings of the 20th ACM International Conference on Multimedia - MM '12*. Don't ask me what i'm like, just watch and listen (ACM Press, New York, 2012), p. 329. doi:10.1145/2393347.2393397. http://dl.acm.org/citation.cfm?id=2393347.2393397. Accessed Sept 2016–Mar 2017
20. J Staiano, B Lepri, R Subramanian, N Sebe, F Pianesi, in *Proceedings of the 19th ACM International Conference on Multimedia - MM '11*. Automatic modeling of personality states in small group interactions (ACM Press, New York, 2011), p. 989. doi:10.1145/2072298.2071920. http://dl.acm.org/citation.cfm?id=2072298.2071920. Accessed Sept 2016–Mar 2017
21. F Celli, F Pianesi, in *Proceedings of the Workshop on Computational Personality Recognition*. Workshop on computational personality recognition (shared task), (2013). http://www.michalkosinski.com/celli_wcpr13.pdf. Accessed Sept 2016–Mar 2017
22. G Chittaranjan, J Blom, D Gatica-Perez, in *2011 15th Annual International Symposium on Wearable Computers*. Who's who with Big-Five: analyzing and classifying personality traits with smartphones (IEEE, 2011), pp. 29–36. doi:10.1109/ISWC.2011.29. http://dl.acm.org/citation.cfm?id=2015035.2015314. Accessed Sept 2016–Mar 2017
23. D Olguin, PA Gloor, AS Pentland, in *Proceedings of the 2009 Aaai Spring Symposium on Human Behavior Modeling*. Capturing individual and group behavior with wearable sensors, (AAAI, California, 2009)
24. N Yee, N Ducheneaut, L Nelson, P Likarish, in *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*. Introverted elves & conscientious gnomes (ACM Press, New York, 2011), p. 753. doi:10.1145/1978942.1979052. http://dl.acm.org/citation.cfm?id=1978942.1979052. Accessed Sept 2016–Mar 2017
25. V Ponce-López, B Chen, M Oliu, C Corneanu, A Clapés, I Guyon, X Baró, HJ Escalante, S Escalera, in *Computer Vision–ECCV 2016 Workshops*. Chalearn lap 2016: First round challenge on first impressions-dataset and results (Springer, Amsterdam, 2016), pp. 400–418

Kindiroglu *et al. EURASIP Journal on Image and Video Processing* (2017) 2017:77

Page 14 of 14

26. HJ Escalante, V Ponce-Lopez, J Wan, MA Riegler, B Chen, A Clapes, S Escalera, I Guyon, X Baro, P Halvorsen, H Müller, M Larson, in *International Conference on Pattern Recognition (ICPR 2016) Workshops*. ChaLearn Joint Contest on multimedia challenges beyond visual analysis: an overview, (Cancun, Mexico, 2016). https://hal.archives-ouvertes.fr/hal-01381144. Accessed Sept 2016–Mar 2017

27. F Gürpınar, H Kaya, AA Salah, in *Computer Vision–ECCV 2016 Workshops*. Combining deep facial and ambient features for first impression estimation (Springer, Amsterdam, 2016), pp. 372–385

28. A Subramaniam, V Patel, A Mishra, P Balasubramanian, A Mittal, in *Computer Vision–ECCV 2016 Workshops*. Bi-modal first impressions recognition using temporally ordered deep audio and stochastic visual features (Springer, Amsterdam, 2016), pp. 337–348

29. B Aydin, AA Kindiroglu, O Aran, L Akarun, in *International Conference on Pattern Recognition (ICPR 2016) Workshops*. Automatic personality prediction from AudioVisual data using random forest regression (IEEE, Cancun, 2016)

30. T Evgeniou, M Pontil, in *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '04*. Regularized multi-task learning (ACM Press, New York, 2004), p. 109. doi:10.1145/1014052.1014067

31. A Evgeniou, M Pontil, in *Conference on Neural Information Processing Systems*. Multi-task feature learning (Curran Associates, Vancouver, 2007). Accessed Sept 2016–Mar 2017

32. A Argyriou, T Evgeniou, M Pontil, Convex multi-task feature learning. Mach. Learn. **73**(3), 243–272 (2007)

33. P Gong, J Ye, C Zhang, in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. Robust multi-task feature learning, vol. 2012 (ACM, 2012), pp. 895–903

34. A Jalali, P Ravikumar, S Sanghavi, C Ruan, in *NIPS*. A dirty model for multi-task learning (Curran Associates, Vancouver, 2010), pp. 1–9

35. Z Kang, K Grauman, F Sha, in *Proceedings of the 28th International Conference on International Conference on Machine Learning*. Learning with whom to share in multi-task feature learning (ICML, USA, 2011)

36. X Chen, S Kim, Q Lin, JG Carbonell, EP Xing, Graph-structured multi-task regression and an efficient optimization method for general fused lasso (2010). arXiv preprint arXiv:1005.3579

37. D Sanchez-Cortes, O Aran, A nonverbal behavior approach to identify emergent leaders in small groups. IEEE Trans. Multimed. **14**(3–2), 816–832 (2012)

38. O Aran, D Gatica-Perez, in *Proceedings of the 15th ACM on International Conference on Multimodal Interaction - ICMI '13*. One of a kind: inferring personality impressions in meetings (ACM Press, New York, 2013), pp. 11–18. doi:10.1145/2522848.2522859

39. O Aran, D Gatica-Perez, in *Proceedings of the International Conference on Multimodal Interaction*. Cross-domain personality prediction: from video blogs to small group meetings, (2013), pp. 127–130. doi:10.1145/2522848.2522858. https://dl.acm.org/citation.cfm?doid=2522848.2522858. Accessed Sept 2016–Mar 2017

40. I Guyon, A Elisseeff, An introduction to variable and feature selection. J. Mach. Learn. Res. **3**, 1157–1182 (2003). doi:10.1162/153244303322.753616

41. R Tibshirani, Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Methodol. **58**(1), 267–288 (1996). doi:10.1007/s13398-014-0173-7.2

42. J Zhou, J Chen, J Ye, MALSAR: Multi-task learning via structural regularization (2011). Technical report, Arizona State University. http://www.public.asu.edu/~jye02/Software/MALSAR. Accessed Sept 2016–Mar 2017

43. J Chen, J Liu, J Ye, Learning incoherent sparse and low-rank patterns from multiple tasks. ACM Trans. Knowl. Disc. Data TKDD. **5**, 4–22 (2012)

44. Y Zhang, DY Yeung, in *Proceedings of the AAAI Conference on Artificial Intelligence*. Multi-task learning in heterogeneous feature spaces, (2011)

45. C-C Chang, C-J Lin, in *ACM Transactions on Intelligent Systems and Technology (TIST)*. LIBSVM: A library for support vector machines, (ACM, New York, 2011), p. 27

46. L Teijeiro-Mosquera, J-I Biel, JL Alba-Castro, D Gatica-Perez, What your face vlogs about: expressions of emotion and Big-Five traits impressions in youtube. IEEE Trans. Affect. Comput. **6**(2), 193–205 (2015)