
Sex Recognition from Faces Using Neural Networks

B. Golomb, T. Sejnowski

*Howard Hughes
Medical Institute
The Salk Institute
10010 North Torrey Pines Road
La Jolla, CA 92037*

1 INTRODUCTION

Recognizing the sex of conspecifics is important: Expending one's attentions Q or one's seed Q on the wrong parity of partner could put one at a competitive disadvantage. While some animals use pheromones to recognize sex, in humans this task is primarily visual: "Many socially living animals ... recognize each other as members of the same species, as individuals, and as social partners by means of visual signals and communicate their mood and intentions by the same sensory modality. In many primate species the individual structure of the face is the most important visual characteristic of each group member" (Grusser, Selke, & Zynda 1985). A core issue is how sex is recognized from face; yet until recently this received little attention.

Many factors make the face a focus of interest in humans, and thus a good locus for sexually dimorphic features. It is the subject of scrutiny in communication: Facial gestures, phylogenetic precursors to verbal communication (Grusser, et al. 1985), supplement verbal language cues; lip reading augments audial cues of language (and may override them, as in the McGurk effect (McGurk, & MacDonald 1976). The human face is visually accessible since it is less clothed for protection and modesty than other bodily parts, to permit vision and breathing. Autonomic and physical signs appear on the face (such as pupillary constriction/dilation, or flared nostrils with labored breathing or

anger), and can indicate the individual's arousal and potential for threat. Too, emotions are played out in facial expressions: These differ from expressions used in language, and may dictate whether approach or avoidance is the best course. (Ekman and Friesen have identified six "universal" expressions of emotion which occur cross-culturally (Ekman 1973a; Ekman 1973b; Ekman 1977; Ekman 1989; Ekman, & Friesen 1969; Ekman, Friesen, & Simons 1985) and have distinctive autonomic nervous system signatures (Ekman, Levenson, & Friesen 1983; Schwartz, Weinberger, & Singer 1981; Sternbach 1962), making them effective indicators of the bearer's state; several of these, such as anger and fear, have decided significance to the onlooker.)

In short, we have reason to look at the face; and we have need of a visual indication of sex; so it is sensible that some visual cues of sex should abide on the face.

Primate physiologists have devoted much attention to faces. Neurons that respond selectively to faces have been identified in the amygdala (Rolls 1981; Rolls 1984; Sanghera, Rolls, & Roper-Hall 1979); the anterior inferior temporal cortex (Gross, Rocha-Miranda, & Bender 1972; Rolls, Judge, & Sanghera 1977); a polysensory area of the dorsal (anterior) bank of the superior temporal sulcus (STS) (Bruce, Desimone, & Gross 1981), where response latencies are 200-300 msec after presentation of the face; in the fundus and anterior portion of the superior temporal sulcus, where response latencies are 70-150 msec (Baylis, Rolls, & Leonard 1985; Perrett, Rolls, & Caan 1982; Perrett, et al. 1984); in parietal cortex (Leinonen, & Nyman 1979); and in frontal cortex (Pigarev, Rizzolatti, & Scandolara 1979). The superior temporal sulcus work is particularly interesting, as up to 20% (Bruce, et al. 1981; Perrett, et al. 1982), though as low as 3% (Perrett, et al. 1984) of neurons tested are accounted "face cells" in that their response to the best tried face was at least twice (and often more than ten times) as great than the greatest response to any nonface stimulus tested (Perrett, et al. 1982) (including gratings and random 3-D objects for instance). Some such neurons appear to identify specific faces independent of expression (Perrett, et al. 1982), while others respond to features or expression independent of face (Perrett, et al. 1984).

There is great variability in how selective these cells are to a specific face: Some are active for most faces, and others highly selective for but one or a few (Perrett, et al. 1982; Rolls 1992). Consequently it would not be implausible for some STS cells to respond preferentially according to sex of faces; or this could occur in the amygdala, where damage leads to disruption

of emotional and social responses to faces (Rolls 1984; Young, & Yamane 1992). No attempt has been made to test for face cells selective for the sex of a primate (human or monkey), perhaps due to lack of conviction that facial sex differences exist in rhesus monkeys.

Monkeys with bilateral temporal lobe lesions (affecting, for instance, the STS) may develop "Kluver-Bucy" syndrome (Horel, Keating, & Misantone 1972; Jones, & Mishkin 1972; Kluver, & Bucy 1939; Trezean, Dalle, & Ore 1955), which includes inappropriate reaction to faces, and sexual indiscriminance. However, this is not likely to ensue from selective inability to recognize sex, since these monkeys' pathological sexual eclecticism extends not only to individuals of the wrong sex, but of the wrong species Q and kingdom.

There may be a human analog of Kluver-Bucy (Trezean, et al. 1955); however the best-described human lapse in responding to faces is "prospagnosia," or inability to recognize faces (Benton, & van Allen 1972; Bodamer 1947; Bornstein 1963; Bornstein, Sroka, & Munitz 1969; Christen, Landis, & Regard 1985; Damasio, Damasio, & van Hoesen 1982; Gloning, Gloning, Jellinger, & Quatember 1970; Hecaen, & Angelergues 1962; Meadows 1974; Whiteley, & Warrington 1977). This attends bilateral mesial occipitotemporal lesions (lesions of the lingual and fusiform gyri) (Damasio, et al. 1982; Jeeves 1984), often from vascular lesions of the posterior cerebral artery. The affected neurons may be the analog of the "face cells" studied in monkeys. The afflicted are able to tell there is a face, but cannot identify whose it is. Sex recognition is typically preserved (Tranel, Damasio, & Damasio 1988).

The lesion technique provides evidence that a region of the brain may be involved in some function, but cannot by itself tell us what that function is. New techniques for functionally mapping the activity of normal brains have recently been developed. One of these methods, positron-emission tomography (PET), has been used to identify brain regions that are involved in face processing and, in particular, to localize brain areas that are active during recognition of sex from faces (Sergent, Ohta, & MacDonald 1992).

Several mechanisms may explain preserved ability to discern sex, in the "face" of lost ability to identify the bearer of the face, in humans. Different cortical areas may subserve recognition of sex versus facial identity; indeed, more rostral areas of temporal cortex, in addition to subcalcarine cortex, are

believed necessary for identity but not sex recognition in humans; damage to these rostral areas would engender loss of identity recognition, with preservation of sex recognition. Loss of ability to discriminate sex appears to require damage to both left and right subcalcarine "early" visual association cortices (Tranel, et al. 1988).

Preserved ability to infer sex from ancillary features may help: Characteristic hairstyles, clothing, jewelry or makeup are associated with males or females, but seldom uniquely identify the bearer. In addition, sex discrimination (but not identity discrimination) is vastly overlearned, in legion contexts, which could promote a robust neuronal representation less sensitive to loss of a fraction of involved neurons. Each face to which we are exposed represents a "training instance" of either maleness or femaleness, with associated features of voice, mannerisms, clothing and body habitus to disambiguate the answer (provide a "teacher"); yet we are exposed to comparatively few instances of any individual, and don't always have a secondary tag to tell us if we are right or wrong about who it is, supposing we have any idea at all.

Quite young infants are able to tell males from females (by whatever constellation of cues, probably including pitch of voice), allowing new encounters to serve as training examples in which novel faces are linked to the appropriate sex. Thus the process of sex recognition training may commence quite early. Finally, computer analysis has revealed that the features which allow sex discrimination may occur in the first few principal components of the faces, involving lower spatial frequency cues, while those for identity recognition reside in later principal components (O'Toole, Abci, Deffenbacher, & Valentin 1993).

In any event, neurologically intact humans do recognize sex from face. But by and large they are unable to say how. Although certain features are nearly pathognomonic for one sex or the other (facial hair for men, makeup or certain hairstyles for women), even in the absence of these cues the determination is made; and even in their presence, other cues may override.

Sex-recognition in faces is thus a prototypical pattern recognition task of the sort at which humans traditionally excel, and by which knowledge-based Artificial Intelligence has traditionally been vexed; in short, an ideal application to demonstrate the capabilities of neural networks. It appears to follow no simple algorithm, and indeed is modifiable according to fashion (makeup, hair etc). While ambiguous cases exist, for which we must appeal to other cues such as physical build (if visible), voice patterns (if audible), and mannerisms, humans are fairly good in most cases at discriminating sex

merely from photos of faces, without resorting to such adscititious cues. The obvious question is, can neural networks do the same? We have developed a neural network system called SexNet that can classify the faces of college-aged students with 89% accuracy. Humans achieved a classification performance of 88% on the same set of faces. Our performance is higher than that reported from other neural network systems that start with feature extraction (Golomb, Lawrence, & Sejnowski 1991). Moreover, even a perceptron, with one layer of variable weights, achieves a respectable accuracy (Gray, Lawrence, Golomb, & Sejnowski 1993). These results demonstrate that good performance at face classification can be achieved with relatively simple preprocessing of images. In this chapter we will summarize these results.

2 METHODS

We started with 512x512 face images of 90 young adult faces (45 male, 45 female) (See Fig. 1) (O'Toole, Millward, & Anderson 1988). Faces had no facial hair, no jewelry, and apparently no makeup. A white cloth was draped about each neck to eliminate possible clothing cues. Most images were head on, but exact angle varied.

Each face image was rotated until eyes were level; scaled and translated to position eyes and mouth similarly in each image; and clipped to present a similar extent of image around eyes and mouth. (This clipping eliminated hair cues on many faces, though residual hair cues remained in some.) Final faces were 30x30 pixels with 12 pixels between the eyes, and 8 pixels from eyes to mouth. The 256 gray-level images were adjusted to the same average brightness. (No attempt was made to equalize higher order statistics.) This served as input to the networks, as shown in Figure 1.

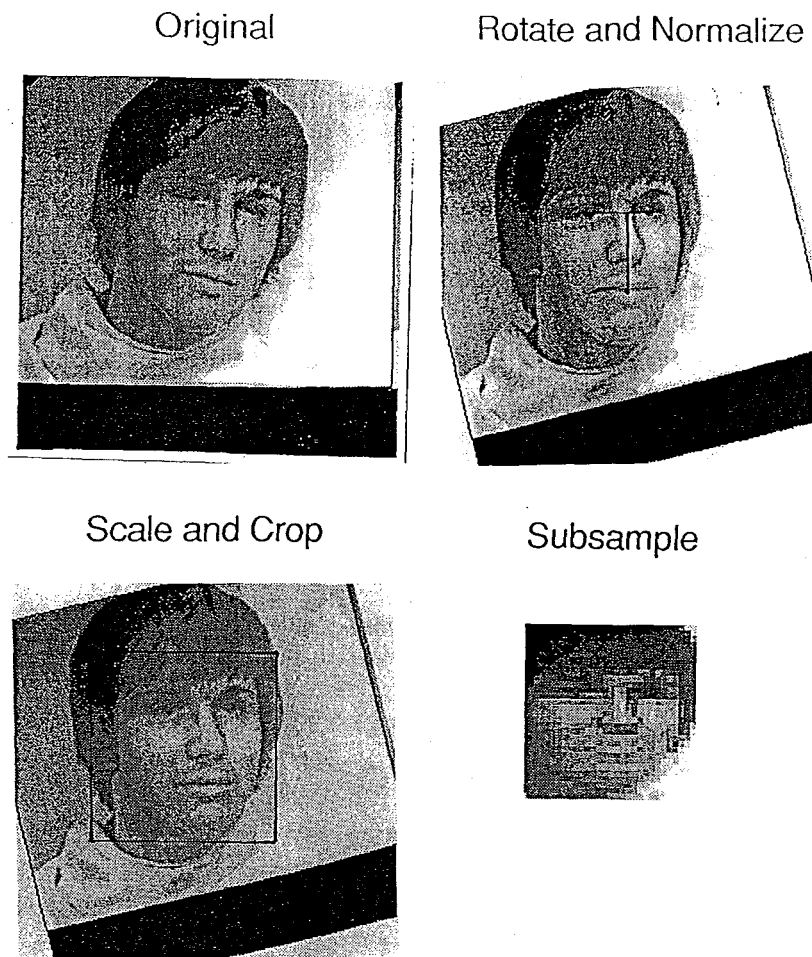


Figure 1

Figure 1: Preprocessing for faces. The original image is shown in the upper left. The center of each eye was located by hand and the line joining the eyes was rotated to the horizontal. The distance between the eyes and the perpendicular distance between the eyes and mouth were normalized (upper right). The image was scaled and cropped (lower left) and subsampled randomly within each subregion to produce a 30x30 image (lower right). The average gray level of the pixels was adjusted to have the same value for all images.

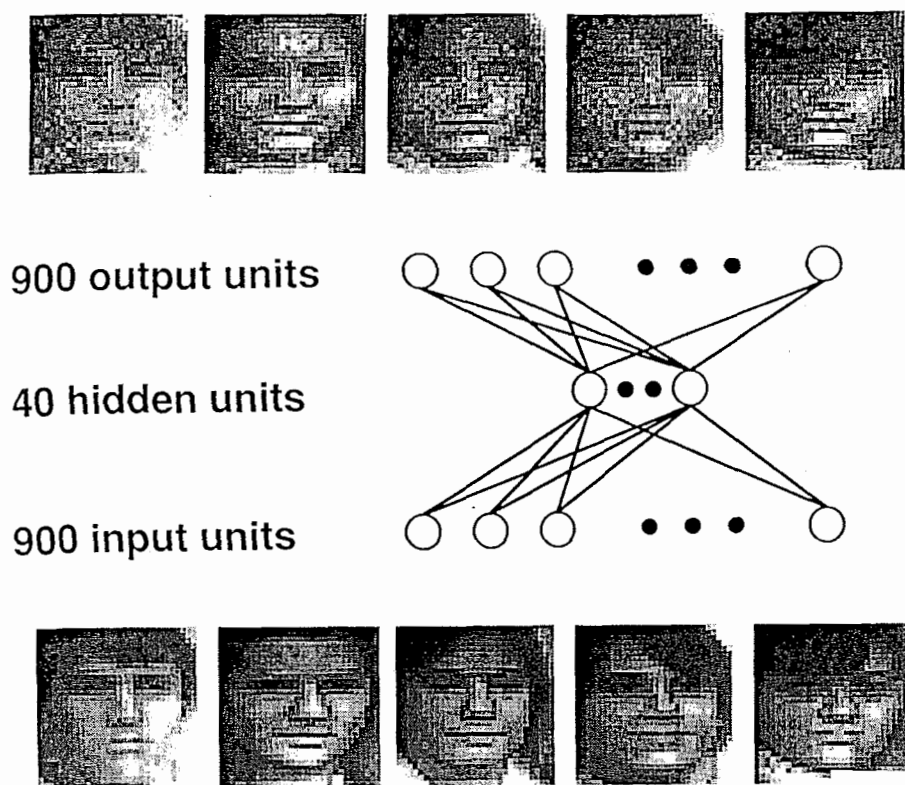


Figure 2

Figure 2: Compression network. The preprocessed images were compressed with a 900x40x900 encoder to reduce the dimensionality of the input space to the SexNet. Sample input images are shown on the bottom and their corresponding outputs from a trained compression network are shown above. In some networks this compression stage was bypassed and the gray level image was used directly as input to the sex identification network.

Network processing consisted of two phases: image compression and sex discrimination. Both networks were fully connected three layer networks with two biases, trained with conventional backpropagation (Parker 1986; Rumelhart, Hinton, & Williams 1986; Werbos 1974), with a bias, a sigmoidal squashing function, and a learning rate of 0.2, using the Neuristique SN2 simulator written by L. Bottou and Y. LeCun.

Image compression followed the scheme of Cottrell, Munro and Zipser (Cottrell, Munro, & Zipser 1987), in which an input to a back-propagation network is reproduced as output, after being forced through a hidden unit bottleneck (i.e. a smaller number of units), as shown in Figure 2. This process, if successful, gives a new reduced representation of the input (here, the face image) in the activity of the (fewer) hidden units. If the network is able to reproduce the input as output from the information in the hidden units, the new reduced representation in the hidden units must contain the salient information of the input (faces). This scheme, though it technically involves supervised learning, can be viewed as unsupervised since no "teacher", other than the input itself, is required to produce the error signal (Cottrell, & Fleming 1990).

The new reduced representation of each input (i.e. the activities of the hidden units of the compression network for that face) can be used as a substitute, more parsimonious input to a second stage network which derives some secondary information from the input (in our case, the sex of the face), as illustrated in Figure 3. The application of this two-stage network technique to faces was pioneered by Cottrell and Fleming, who used it for face recognition, with a more limited effort at sex recognition (producing 37% errors) (Cottrell, et al. 1990).

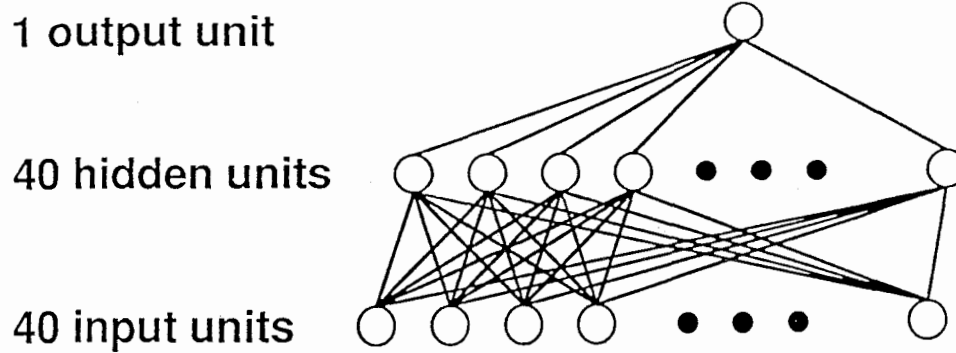
In our work, the compression net served to force the 900 unit (30x30) images (900 inputs) through a 40 hidden unit bottleneck, and reconstruct the image at the 900 unit output level. Thus, the input equalled the desired output. The network trained for 2000 runs on each of 90 faces, yielding output faces which were subjectively distinct and discriminable, although not identical to the inputs. This procedure served to forge a representation of each face in the activities of only 40 units, and thus provide a more tractable input (40 units rather than 900) to the sex discrimination network. A boon for a small training set. The second, sex-discrimination network, had 40 inputs (the activities of the 40 hidden units of the compression net), 0 or 40 hidden units, and one output unit. Training consisted of encouraging, by gradient descent (Rumelhart, et al. 1986) the network to produce a "1" for men, and a "0" for

women. (These numbers are standard network integers, and any similarity to discriminating male and female anatomical parts is purely coincidental.) Values greater than 0.5 were accounted "male", and those less than 0.5 female. Trials were also done with no antecedent compression, and with no hidden units in the SexNet (i.e. feeding the 900 unit input directly to the one unit output).

As the suitable benchmark for sex discrimination by the network is human performance on the same faces, human testing was undertaken. Sex order was chosen to be random for 45 faces (even vs odd sequential digits of π coded male vs female), and, to equalize the numbers of males and females, the order was repeated with reverse parity for the second 45 faces. 5 humans were tested on the 90 faces, and made two binary decisions for each face: sex and certainty of their answer (sure vs unsure). They had unlimited time, and could scrutinize faces in any manner. The full 512x512 resolution images were used, which included hair but no clothing (see Fig. 1).

For comparison, 9 tests of the SexNet were undertaken, each training on a different 80 faces, leaving a distinct set of 10 untrained faces for testing. We performed tests on two sets of images. The first set included some hair (See Fig 2) but the second set was cropped so that only the central portion of the face was visible.

Sex Network



Inputs = Compressed Representation
Output = Sex
1 = male (>0.5)
0 = female (<0.5)

Figure 3: Network for identifying sex from compressed images of faces. the 40 hidden units from a compression network (see Fig. 2) are used as input to a 40 x 40 x 1 identification network. The output was thresholded to classify each image as male (output > 0.5) or female (output < 0.5).

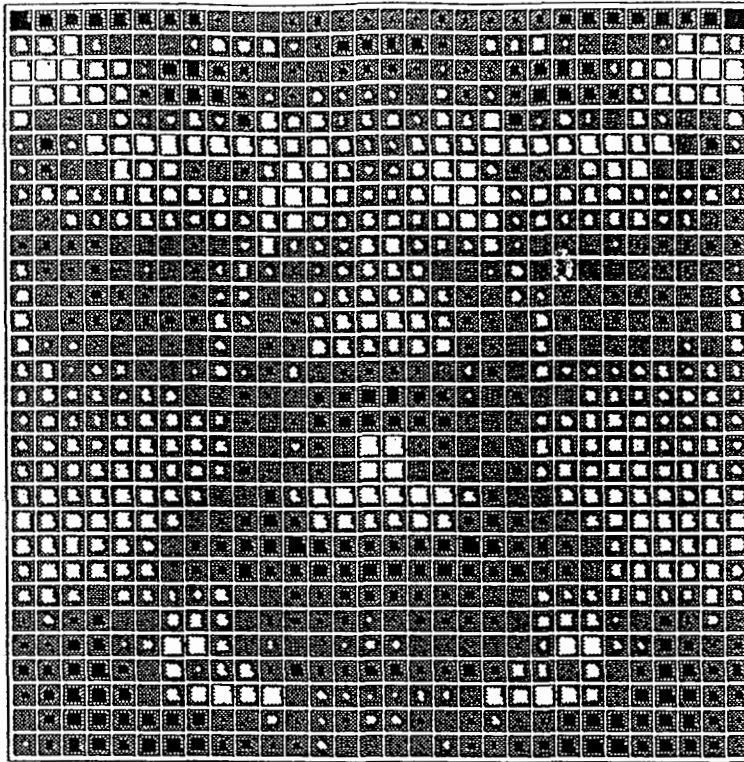


Figure 4: Weights in a perceptron network that was trained to discriminate sex from images of faces. Only the central region of the face was used in training, which excluded information from the outline of the face and hairline. The area of each square in the diagram is proportional to the value of the weight in the network assigned to that region of the face. White weights are excitatory and black weights are inhibitory. A high intensity in a region of the face with a white weight provides evidence for male, and a high intensity region with black weights is evidence for female (Adapted from (Gray, et al. 1993)).

3 RESULTS

Human Performance.

Psychophysical studies of 5 humans on the 90 faces revealed errors on 8, 10, 12, 8, and 14 faces, corresponding to 8.9%, 11.1%, 13.3%, 8.9% and 15.5%, for an average error rate of 11.6%. Humans and the network seemed prone to similar errors. One male face gave particular trouble to the SexNet, being wrongly assigned when a test face, and taking long to train when a training face. This same face was (erroneously) judged "female", "sure" by all 5 observers.

SexNet.

SexNet achieved a performance of 89% on the set of images that included the outline of the face and some hair, which was comparable to the 88% performance achieved by humans given the same information. When the images of the faces were pared down to eliminate the chin and forehead (which also removed all hair cues), sex judgements were much more difficult for human observers; the performance of the network fell from 89% to 81%, and remained superior to human observers.

The hidden units proved unnecessary to the SexNet's performance; a perceptron (no hidden units) was as successful as networks with up to 40 hidden units, suggesting that the task may be linearly separable. In a followup study (Gray, et al. 1993) we trained a perceptron directly from the gray level images and achieved a performance of 79% for the cropped images. This was quite surprising since it had been previously thought that sex discrimination was a higher-order problem. In Figure 4, the area of each square corresponds to the strength of the weight connecting that image pixel to the output unit. The value of the output unit was computed by multiplying the gray level of each pixel with its corresponding weight, summing the products, and passing the sum through a sigmoid. Since "male" corresponded to an output value of 1, each positive term in the sum (from a positive weight) could be considered evidence that the face is male, and each negative term (from a negative weight) evidence that the face is female.

Evaluation of optimal resizing of the faces (prior to their functioning as inputs to the network) showed the choice of 30x30 images to have been near optimal for network performance (reference). In contrast humans need higher resolution for best performance.

Antecedent compression of faces was not required for good performance of the SexNet, which functioned approximately equally well with the raw gray-level images. (However, having once compressed the images, training time was faster with the reduced input due to fewer connections in the network.)

Although compression was not required for good network performance, use of compression permitted examination of the characteristics of hidden units, which in turn provided insight into characteristics used by the network. (Certainly, the characteristics used by that network with those hidden units; and presumably networks without.) If one hidden unit in the compression subnet is "turned on" (activity set to 1), with all others "off" (activity 0), the compression network's output (the decompressed image) reveals a rather vague-appearing nonspecific face, corresponding to the "holon" of Cottrell and Fleming (Cottrell, et al. 1990), or a mixture of the "eigenfaces" of Turk and Pentland (Turk, & Pentland 1991).

Deciphering characteristics of the SexNet was also possible by examining weights in the network: For example, evaluation of the weights of the SexNet showed presence of an elongated philtrum (the space between the nose and mouth, relative to the distance from eyes to mouth which was standardized) was strongly correlated with maleness. This was apparent both in the simplest perceptron networks (See Fig. 4) and in examination of the hidden units from compression inputs with a particularly strong positive connection to maleness, and a strong negative connection to femaleness.

4 DISCUSSION

Comparison with human performance.

The complex visual pattern recognition task of sexing human faces can be adequately performed by a neural network without prior feature selection and with minimal preprocessing Q contrary to the confident predictions of several of our colleagues. Human performance was matched by a using either a 900 x 40 x 900 back-propagation image compression network such as had previously been used by Cottrell and Fleming for face identification, the activities of whose hidden units served as inputs to a 40x40x1 back-propagation SexNet; or equivalent performance was achieved by a perceptron SexNet which translated the 900 unit input (or the 40 unit reduced input) directly to sex with no intermediate hidden units, indicating that the task is

linearly separable. No efforts to optimize the network were needed to match human performance¹.

The SexNet performance was similar to humans' not just by percent errors. Not only did it correctly sex previously unseen faces, but it had difficulties on faces which also posed difficulties for humans. Indeed on one preliminary trial it correctly assigned all ten test faces, but misjudged two of the 80 training faces. These included the problematic male discussed above, to which it assigned the androgynous value of 0.495, and another male on which it performed wretchedly, with a value between 0.2 and 0.3, despite copious training. The SexNet proved correct: The face was an unambiguous female whose sex value had been mistranscribed in the training data. The SexNet correctly sexed the face based on the other faces, in spite of faulty training information. It had evidently done a fine job of abstracting what distinguishes the sexes.

Failure of humans and the network on the same face suggests how one might handle the network's difficulty, in analogy with human strategies. When a face is found by the network to be female (according to weights which correctly gauge sex of "most" training faces) but the person is male (or vice versa), one shouldn't emend male-female categories too drastically, as one may encounter another nearly identical face which is in fact female. The human strategy confronted with a "training face" (one for which sex is known by other criteria) would consist in making a special category for the individual, and having that provide input to overrule the facial information. If another network were trained to identify problematic individuals then they could be dealt with separately. The outliers could then be correctly identified without adverse consequences to generalization. A similar strategy has proved useful in other domains: In Baxt's neural network which identifies whether patients presenting to the emergency room with chest pain are having a myocardial infarction, superior performance was achieved by training a separate network on the difficult-to-learn cases (Baxt 1992).

¹ The SexNet was in fact slightly better at discriminating sex than the humans tested; however, since the network's performance improved with training, the results may be taken to indicate that the humans tested should spend more time engaged in discriminating sex.

Comparison with other network approaches to sex recognition.

Others have used neural networks for sex identification from faces. The task is referred to by other authors as "gender" recognition, but "gender" is properly a grammatical term: specifically, words have gender, people don't. In English, the term "gender" has attained increasing currency as a euphemistic substitute for (male vs female) "sex". However, "To talk of persons or creatures of the masculine or feminine g., meaning of the male or female sex, is either a jocularity (permissible or not according to context) or a blunder."² Admittedly, the usage is otherwise, but we prefer the anatomically correct term. Neural network efforts on sex recognition begin with Cottrell and Fleming, whose principal focus was face recognition. Their network used a 64 x 64 face with 80 hidden units for compression; this was passed to a perceptron with a unit for each name, a unit for faceness, and units for male and female "gender". They suggest that for image compression the relevant ratio is not that of hidden units to input units, but rather of hidden units to number of patterns, which they suggest should be about 1:1. (Ours was 1: >2.). Their training set consisted of 64 faces and 13 non-face stimuli. They note that their network "failed to accurately categorize novel faces according to gender, making 37% errors on novel faces"; however their the small training set is likely to contribute.

Brunelli and Poggio used Hyper Basis Function networks for "gender" classification. Their strategy differed from ours and that of Cottrell in that they first automatically extracted features from images of pre-normalized scale and rotation; these features, rather than the faces, were used as input. This requires that they have advance insight into which features might be important for the discrimination process; this is a task which is bypassed in the strategy we employed, which requires no advance insight into what might prove useful. They also symmetrized the faces by averaging left and right eyebrow and chin information, which may be reasonable for computer sexing, but detracts from any suggestion of physiological verisimilitude. They used two competing network, one for male recognition and one for female recognition; the outcome was determined by which network gave the greatest response. Only three of the sixteen features used as input developed

² Fowler's Modern English Usage, Oxford University Press, 2nd ed. 1985, revised 1983

significant weights, and two of these may be artifacts of modern sex-specific preening practices: eyebrow thickness and distance from eyebrows to eyes are both artificially modified in women, in the direction of increased eye-to-brow distance and decreased width, by the practice of eyebrow plucking or tweezing (which may be even more widespread in Italy, whence the face images derived, than in the U.S.), so that their preeminence among the discriminating features may rank with hair style, makeup, and jewelry as social rather than biological facial cues. The third feature was nose width. Vertical position of nose and mouth were also given small weights. Their strategy of using competing male and female networks allowed natural determination of the networks' "prototype" male and female faces, constrained, however, by the features selected for the network. These prototypes were observed to resemble not average faces, but caricatures of the sexes, exaggerating the distinctive features. Interestingly, the prototype faces extracted from the networks, in addition to differing in eyebrow findings (and face size) differed perhaps most markedly in the size of the philtrum (which had been identified as important to the SexNet)! Indeed, the nose width, the third "important" feature, was not visibly different in the prototypes. However female nose length is represented as much longer than male, which is introspectively unrealistic Q but may have resulted from a drive to display a small difference between eye-nose and eye-mouth distances (small philtrum). This reinforces the concern that the features selected for use in a feature-driven network may not be the most important (philtrum span was not among the input features mentioned, though it can be obtained by differencing two other features). Additionally, the much longer nose length for the female than the male prototype suggests that prototype faces driven by preselected features cannot be taken at "face" value. This feature-based network showed an average performance of 79% for sexing new faces, compared to 89% for the SexNet including an outline of the face and 81% including only the central region of the face.

Other applications of SexNet.

Although the SexNet task has limited utility of itself Q after all, humans sex human faces fine, without recourse to a neural network Q extensions of this work have application. For instance, it is not known whether faces differ for male and female rhesus monkeys. By training a neural network to perform monkey sexing, and comparing the network's performance on untrained monkey faces to their known sex, better than chance performance would imply there are facial sex differences in rhesus monkeys Q answering a

question of some ethological significance. Physiologists might then be inclined to look for sex-specific face cells in monkey cortex.

In a more lighthearted vein, one could use personality indices rather than sex for the second phase of the net, to scientifically test the tenets of anthroposcopy (physiognomy), possibly for the first time.

More importantly a variety of congenital medical disorders (such as Down syndrome) are accompanied by "craniofacial anomalies" (Dyken, & Miller 1980), resulting in distinctive "facies", or facial appearances. Some are subtle and/or rare, and not often recognized by physicians. It may be possible to screen normal from affected infants or children using special purpose neural networks. We hope to extend our work to encompass neural nets for diagnosing William's syndrome, or infantile hypercalcemia, in which children's faces are "elfin-like" (Bellugi, Bihle, Trauner, Jernigan, & Doherty 1990; Trauner, Bellugi, & Chase 1989). Williams' faces compare to normals in a manner which recalls the male / female distinction in that no isolated well described features occur in all of one but none of the other; rather the gestalt distinguishes them. Early diagnosis is important because these children often have associated cardiac defects requiring surgical correction.

We have extended our work to include neural network analysis of facial expressions, and using similar methodology to that described above, have successfully achieved discrimination among eight facial expressions (with examples of each expression differing in intensity) corresponding crudely to smile, frown, brow-raise, purse-lips, pucker-lips, sneer, squint, and neutral (unpublished). The neutral expression was the most difficult to train, a finding which again correlated to that of human observers, who like the network failed to misclassify this expression as any other, but were reluctant to denote it neutral, either.

References

- Baxt, W. G. (1992). Improving the accuracy of an artificial neural network using multiple differently trained networks. *Neural Computation*, 4, 772-780.
- Baylis, G. C., Rolls, E. T., & Leonard, C. M. (1985). Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Research*, 342, 91-102.
- Bellugi, U., Bihle, A., Trauner, D., Jernigan, T., & Doherty, S. (1990). Neuropsychological, neurological, and neuroanatomical profile of Williams syndrome children. *American Journal of Medical Genetics*, In Press.
- Benton, A. L., & van Allen, M. W. (1972). Prosopagnosia and facial discrimination. *Journal of Neurological Science*, 15, 157-172.
- Bodamer, J. (1947). Die Prosopagnosie. *Archiv für Psychiatrie und Nervenkrankheiten*, 179, 6-53.
- Bornstein, B. (1963). Prosopagnosia. In L. Halpern (Ed.), *Problems of dynamic neurology* New York: Grune and Stratton.
- Bornstein, B., Sroka, H., & Munitz, H. (1969). Prosopagnosia with animal face agnosia. *Cortex*, 5, 164-169.
- Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, 46, 369-384.
- Christen, L., Landis, T., & Regard, M. (1985). Left hemispheric functional compensation in prosopagnosia. A tachistoscopic study with unilaterally lesioned patients. *Human Neurobiology*, 3 or 4,
- Cottrell, G., & Fleming, M. (1990). Face recognition using unsupervised feature extraction. Paris, France: Kluwer Academic Publishers, 322-325.
- Cottrell, G., Munro, P., & Zipser, D. (1987). Learning internal representations of gray scale images: An example of extensional programming. Seattle, Wa.:
- Damasio, A. R., Damasio, H., & van Hoesen, G. W. (1982). Prosopagnosia: anatomic basis and neurobehavioral mechanisms. *Neurology*, 32, 331-341.

Dyken, P. R., & Miller, M. D. (1980). *Facial Features of Neurologic Syndromes*. St. Louis, Missouri: C.V. Mosby Company.

Ekman, P. (1973a). Cross-cultural studies of facial expression. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review* (pp. 169-222). New York: Academic Press.

Ekman, P. (1973b). *Darwin and Facial Expression: A Century of Research in Review*. New York: Academic Press,

Ekman, P. (1977). Biological and cultural contributions to body and facial movement. In J. Blacking (Ed.), *Anthropology of the Body* (pp. 39-84). London: Academic Press.

Ekman, P. (1989). The argument and evidence about universals in facial expressions of emotion. In H. W. a. J. Manstead (Ed.), *Handbook of psychophysiology: Emotion and social behavior* (pp. 143-164). London: John Wiley and Sons.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, 32(1), 88-105.

Ekman, P., Friesen, W. V., & Simons, R. C. (1985). Is the startle reaction an emotion? *Journal of Personality and Social Psychology*, 49(5), 1416-1426.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes between emotions. *Science*, 221, 1208-1210.

Gloning, I., Gloning, K., Jellinger, K., & Quatember, R. (1970). A case of "prosopagnosia" with necropsy findings. *Neuropsychologia*, 8, 199-204.

Golomb, B. A., Lawrence, D. T., & Sejnowski, T. J. (1991). SEXnet: A neural network identifies sex from human faces. In D. S. Touretzky, & R. Lippmann (Ed.), *Advances in Neural Information Processing Systems*, Vol. 3 San Mateo, California: Morgan Kaufmann.

Gray, M., Lawrence, D., Golomb, B. A., & Sejnowski, T. J. (1993). A perceptron reveals the face of sex. UCSD Institute for Neural Computation Technical Report INC 93-03.

- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Neurophysiology*, 35, 96-111.
- Grusser, O. J., Selke, T., & Zynda, B. (1985). Age dependent recognition of faces and vases in children and adolescents. *Human Neurobiology*, 4, 33-39.
- Hecaen, H., & Angelergues, R. (1962). Agnosia for faces (prosopagnosia). *Archives of Neurology*, 7, 92-100.
- Horel, J. A., Keating, E. G., & Misantone, L. G. (1972). Kluver-Bucy syndrome produced by destroying neocortex or amygdala. *Brain Research*, 94, 347-359.
- Jeeves, M. A. (1984). The historical roots and recurring issues of neurobiological studies of face perception. *Human Neurobiology*, 3, 191-196.
- Jones, B., & Mishkin, M. (1972). Limbic lesions and the problem of stimulus-reinforcement associations. *Experimental Neurology*, 36, 362-377.
- Kluver, H., & Bucy, P. C. (1939). Preliminary analysis of functions of the temporal lobes in monkeys. *Arch. Neurol. Psychiat.*, 42, 979-1000.
- Leinonen, L., & Nyman, G. (1979). Functional properties of cells in anterolateral part of area 7 associative face area of awake monkeys. *Experimental Brain Research*, 34, 321-333.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
- Meadows, J. C. (1974). The anatomical basis of prosopagnosia. *J. Neurol. Neurosurg. Psychiat.*, 37, 489-501.
- O'Toole, A. J., Abci, H., Deffenbacher, K. A., & Valentin, D. (1993). Low-dimensional representation of faces in higher dimensions of the face space. *Journal of the Optical Society of America*, A10, 405-411.
- O'Toole, A. J., Millward, R. B., & Anderson, J. A. (1988). A physical system approach to recognition memory for spatially transformed faces. *Neural Networks*, 1, 179-199.

Parker, D. B. (1986). A comparison of algorithms for neuron-like cells. In J. S. Denker (Ed.), *Neural networks for computing* New York: American Institute of Physics.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47, 329-342.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology*, 3, 197-208.

Pigarev, I. N., Rizzolatti, G., & Scandolara, C. (1979). Neurones responding to visual stimuli in the frontal lobe of macaque monkeys. *Neuroscience Letters*, 12, 207-212.

Rolls, E. T. (1981). Responses of amygdaloid neurons in the primate. In Y. Ben-Ari (Ed.), *The Amygdaloid Complex* (pp. 383-393). Amsterdam: Elsevier.

Rolls, E. T. (1984). Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Human Neurobiology*, 3, 209-222.

Rolls, E. T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philosophical Transactions of the Royal Society of London*, B335, 11-20.

Rolls, W. T., Judge, S. J., & Sanghera, M. K. (1977). Activity of neurones in the inferotemporal cortex of the alert monkey. *Brain Research*, 130, 229-238.

Rumelhart, D. E., Hinton, G., & Williams, R. J. (1986). Learning internal representation by error propagation. In D. E. R. a. J. L. McClelland (Ed.), *Parallel Distributed Processing, Explorations in the microstructure of cognition* (pp. 318-362). Cambridge, Mass.: MIT Press.

Sanghera, M. K., Rolls, E. T., & Roper-Hall, A. (1979). Visual responses of neurons in the dorsolateral amygdala of the alert monkey. *Experimental Neurology*, 63, 610-626.

Schwartz, G. E., Weinberger, D. A., & Singer, J. A. (1981). Cardiovascular differentiation of happiness, sadness, anger and fear following imagery and exercise. *Psychosomatic Medicine*, 43, 343-363.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain*, 115, 15-36.

Sternbach, R. A. (1962). Assessing differential autonomic patterns in emotions. *Journal of Psychosomatic Research*, 6, 87-91.

Tranel, D., Damasio, A. R., & Damasio, H. (1988). Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology*, 38(5), 690-696.

Trauner, D., Bellugi, U., & Chase, C. (1989). Neurologic features of Williams and Down Syndromes. *Pediatric Neurology*, 5(3), 166-168.

Trezean, H., Dalle, & Ore, G. (1955). Syndrome of Kluver and Bucy reproduced in man by bilateral removal of the temporal lobes. *Neurology*, 5, 373-380.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71-86.

Werbos, P. (1974). *Beyond Regression: New tools for prediction and analysis in the behavioral sciences*. Harvard University,

Whiteley, A. M., & Warrington, E. K. (1977). Prosopagnosia: a clinical, psychological and anatomical study in three patients. *J. Neurol. Neurosurg. Psychiat.*, 40, 394-430.

Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science*, 256, 1327-1331.