Report-2: MUZZLE ANALYSIS for BIOMETRIC IDENTIFICATION and UNSUPERVISED CLUSTERING RESULTS

- ***Dr. Kannan Karthik (EEE Dept., IIT Guwahati, June 26, 2017)***

Motivation and Executive Summary

The report is not just directed towards the search for biometric features in the MUZZLE or SKIN domain of PIGS, but also throws open certain conjectures regarding the potential use of certain patterns present in the MUZZLE towards constituting a BIOMETRIC IDENTIFIER. The report addresses the following questions:

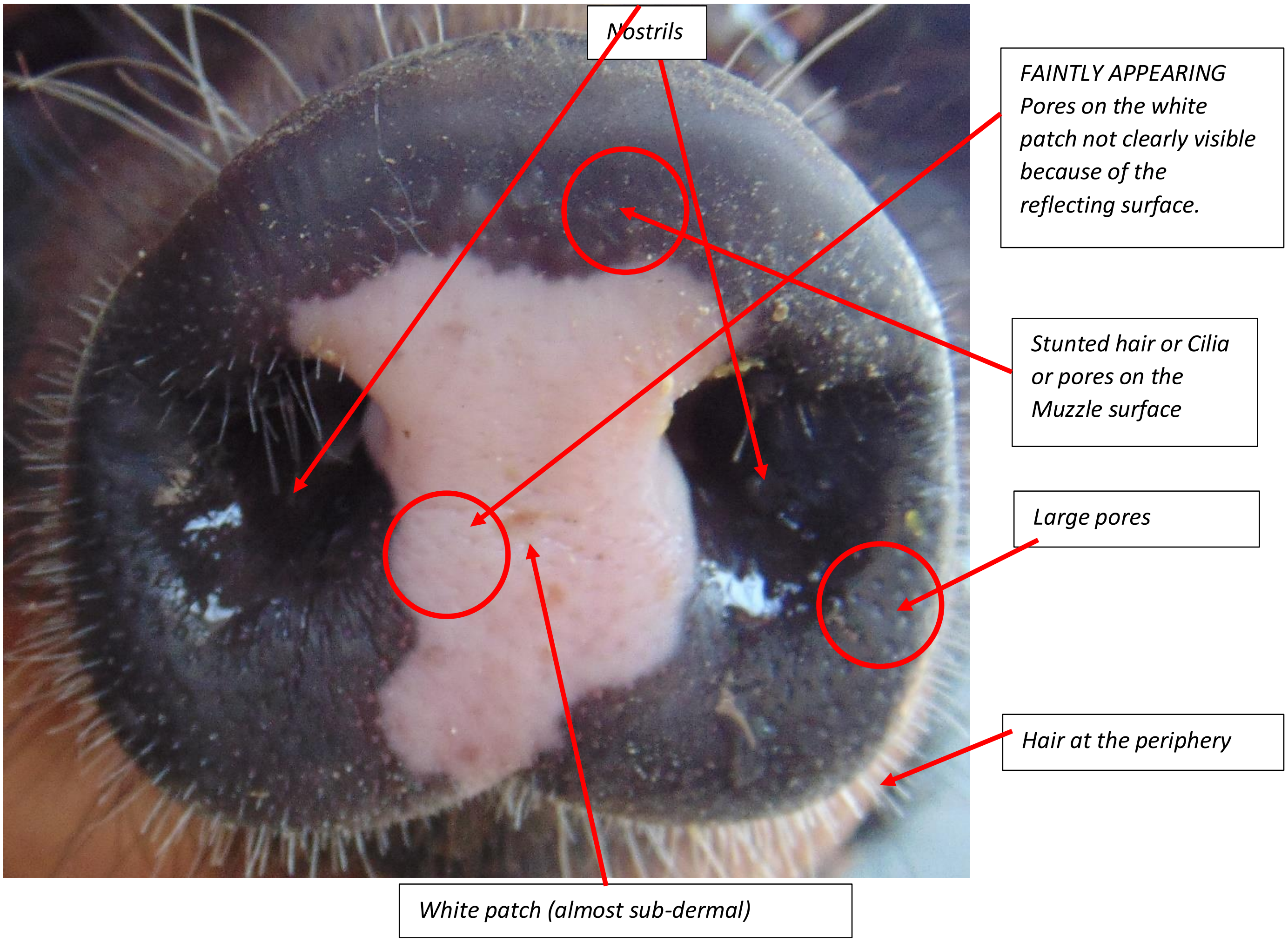
- What are the prominent and stable patterns visible in the MUZZLE of the pig?
- Is it possible to amplify certain features or patterns which are not apparent in the visual domain (i.e. the color images captured)? What information do these hidden features carry? Will these features remain stable over time? Are these features distinctive? Can they be used to derive a biometric identifier? For example the cilia and pores on the surface of the pig form one set of hidden features, the exact role of which is not completely clear. If these are likely to remain stable over time they can be used ANALOGOUS TO MINUTIAE POINTS on a fingerprint.
- Can MUZZLE coloration profiles be used for breed identification or at least shortlisting breeds? A part of the report has been invested in skin patch analysis but is not entirely conclusive.

An algorithm has been developed to amplify certain hidden features in the MUZZLE by constructing a GRADIENT SIGNIFICANCE MAP and the corresponding patch statistic profile. Given the small size of the dataset (only 4 pigs with about 15 variations each), an UNSUPERVISED CLUSTERING algorithm based on k-means was used to segregate a random mixture of MUZZLE IMAGES into these four subject classes. Results close to 100% were obtained for this small dataset for specific parameter settings (patch size vs image size equivalent to 1:10, smoothing parameter $\sigma = 2$ and normalized gradient threshold$\delta_G = 1$).

Introduction

The snout of the pig plays a pivotal role not just in natural breathing, but also for sniffing, understanding and assessing the environment. With some cognitive assistance, it is possible for the pigs to sniff out and hunt preys, distinguish between various varieties of food and sense any foreign presence in the neighborhood. The front portion of the “nose” of the pig is called the muzzle and the two slits which are used both for breathing and smelling are called the “nostrils”. The following elements are visible on or around the muzzle (i.e. at its periphery):

1) The muzzle appears as a contorted ellipse and hosts two nostrils, plenty of stunted hair, hair follicles and pores. There may or may not be a white patch on the muzzle depending on the breed of the pig.
2) Hair or Cilia are present at the periphery of the muzzle. They are also present throughout the muzzle on the inner side. The white patch almost seems like a sub-dermal extension of the layer beneath and does not contain any hair or cilia. However the density of pores remains the same on the white patch. This is naturally inherited in the case of the Hampshire also in some cross-breeds involving Hampshire. The role of this white patch is not clear.



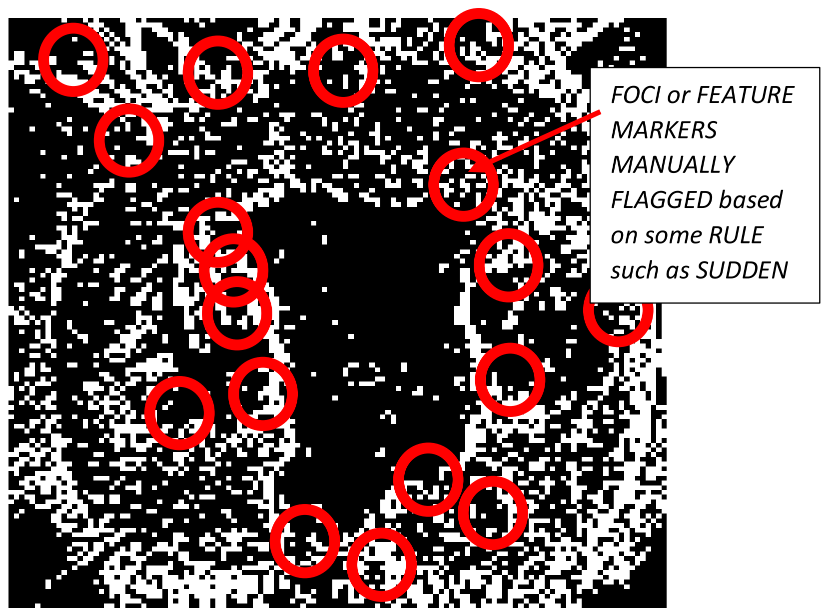
3) The nostrils which continue as the nasal passage within the snout of the pig may form a key feature for biometric identification. The internal curvature of the nostril and the passage could be harnessed from the partially specular/reflecting surface. Specularity of any glossy surface is a function of the geometry and the position of the light source. This can very well be reversed to gauge the shape and orientation of the surface provided the internal details within the nostrils are amplified.

In this report we open up this problem on multiple fronts by jumping ahead a few steps. More questions, conjectures, results and analysis will follow:

1) The shape and orientation of the white patch (visible in Hampshire and selective cross-breeds of Hampshire) is DISTINCT to the animal subject.
2) The position of the nostrils (separation, size and location with respect to the outer envelope varies slightly from subject to subject).

3) The ciliary patterns (and pores) are visible and apparent over the black region but are virtually masked in the white region. This differential appearance (despite having the same density) can be used to re-affirm the white patch detection process. We will see later that because of the specularity of the white surface on the muzzle, the gradient profiling algorithm does not pick up the pores on the white patch, creating a non-uniformity in the gradient significance maps.

4) The locations of the hair follicles or cilia is likely to remain stable for most adult pigs. It is quite possible that old pores do not disappear but rather newer ones may appear. The locations of these DOTS in the gradient significance map shown in Fig. 3 could be used analogously as MINUTIAE for a fingerprint. With a suitable trimming procedure it may be possible to select biometrically significant points (located at these pores) and leave out the irrelevant ones. It may be possible to pivot around groups of points to create FOCI or FEATURE MARKERS which can be used not just for registering the MUZZLE but can also be used for matching two different feature sets. For example in locations where the density of pores is very high, the point can be flagged or marked as a significant feature. Even corner points and bends can be flagged.



5) Even the envelope of the muzzle pattern (or the shape contour) of the muzzle varies slightly from subject to subject.
6) The coloration of the muzzle is also distinct not just because of the presence of pinkish red patches on some of the pigs but is also expected to diversity into different shades of blackish-brown for different pigs. It is anticipated that seat of all colors is BROWN and by adding black or white to brown lighter or darker shades of brown can be constructed. This process is termed as color synthesis which is discussed in the next section.
7) The earlier points can be combined to generate a UNIQUE BIOMETRIC IDENTIFIER.

Layout of the report

Section-1) PATCH COLORATION ANALYSIS and OPEN QUESTIONS

Begin with the inheritance of white patches in the case of Hampshire pigs and raise a few questions regarding the stability of these patches over time. We also examine the body skin coloration for Hampshire and Ghungroo pigs in search of a distinguishing feature but skin color alone may not be sufficient. Body contours and other features may be required.

Section-2) HAIR/CILIARY FEATURES along with PORES on the MUZZLE

We return to the MUZZLE frame and find a way to amplify the "pore and ciliary" structure on the muzzle surface through a gradient profile. We then generate patch statistics (secondary feature) to produce necessary abstractions of the muzzle profile. These statistical abstractions are then fed to an UNSUPERVISED CLUSTERING algorithm to examine its validity on two fronts: 1) Robustness to camera positional changes and background changes (Intra class robustness); 2) Sensitivity or ability to discriminate muzzle shots taken from different pigs (may or may not be from the same breed).

Section-3) UNSUPERVISED CLUSTERING SETUP and RESULTS

Since the data set was small, the best way to validate the feature is to select random mixed images from all classes and then classify them in an UNSUPERVISED fashion. Results were very promising for the GRADIENT PROFILE.

Section-1) PATCH COLORATION ANALYSIS and OPEN QUESTIONS

a) Pig: 41092

b) Pig: 41093

c) *Pig: 41097*

d) Pig: 41099

White patch inheritance conjecture:

There is a conjecture that this white patch inheritance visible in Hampshire and Cross-breeds is largely hereditary with less of a connection with environmental conditions and eating/behavioral habits. The analogy in the case of human beings is the base complexion of the skin which is primarily inherited but undergoes some modifications based on the climate and behavioral habits: viz. An outdoor worker may be fair in base appearance but ends up with a darker skin mainly due to his/her external routines where he/she is exposed more frequently to the sun.

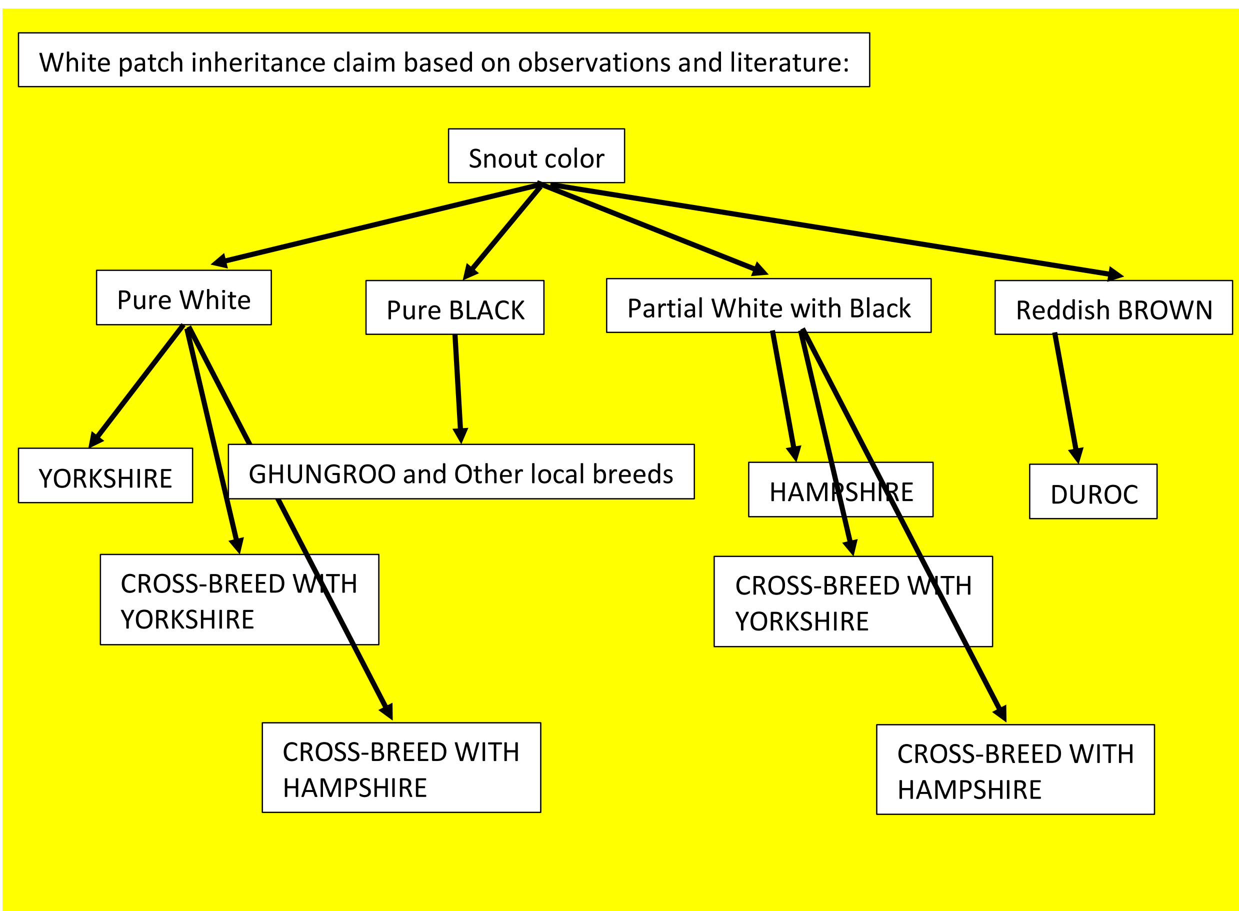
Stability of the white patch over time:

There is a question whether the relative areas of white and dark patches will remain fixed for a specific pig (a or b or c). This STABLITY will hold if and only if the white patches are purely hereditary which implies that the ancestral profile and the dominance of certain genes will control the prominence of the white patch in the progeny.

Variability across siblings, cross-breeds and their progeny etc.

HAMPSHIRE is a pure breed characterized by a white band (either partial or full) near the base of the neck. The width of this band may vary across siblings and within the species. However in the case of CROSS-BREEDs such as those obtained by COMBINING the GHUNGROO and the HAMPSHIRE, the inheritance of the belt is unpredictable. Some piglets from this cross-breed combination may be completely black. A part of this white patch coloration is reflected in the snout. Only HAMPSHIRE and CROSS-BREEDS of HAMPSHIRE and YORKSHIRE (with GHUNGROO) may inherit white patches in the snout. Note that the Yorkshire is completely WHITE.

Fig. 2: Coloration inheritance with cross-bred piglets [Cross-breeding between Ghungroo and one of the other exotic breeds such as Hampshire (or Yorkshire)]. Some piglets have white bands, some piglets are completely black while some are completely white (un-predictability of the white patch inheritance profile).



BODY SKIN PATCH COLORATION ANALYSIS OF PIGS [same analysis can be extended to the muzzle profiles]

Unlike HUMAN SKIN which is set in shades of BROWN owing to the presence of MELANIN, the PIG's skin is set in shades from white for the Yorkshire all the way to black for the Ghungroo. The RED, GREEN and BLUE channels have virtually the same intensity level over the PIG's skin patches. Some exemplar patches of various varieties of GHUNGROO and HAMPSHIRE are shown here.

Ghungroo 21784-female

Ghungroo 21799-male

Ghungroo 21800-male

Ghungroo 21802-female

Hampshire female 41113

Hampshire male 41114

Hampshire male-41093

The mean and variance in the coloration profile for each of the following patches is shown below:

0.36	0.05	0.21	0.07	0.23	0.09	0.18	0.07	0.3	0.08	0.51	0.23	0.86	0.09	0.35	0.06
0.31	0.05	0.2	0.07	0.23	0.09	0.18	0.07	0.28	0.07	0.48	0.22	0.77	0.07	0.33	0.07
0.35	0.05	0.21	0.07	0.25	0.11	0.2	0.09	0.33	0.09	0.49	0.21	0.74	0.08	0.38	0.08

Corresponding to each patch there are six numbers: means and variances of the RED, GREEN and BLUE channels. Observe from these numbers that,

$$\mu_R \approx \mu_G \approx \mu_B = \mu_{patch(i)}$$

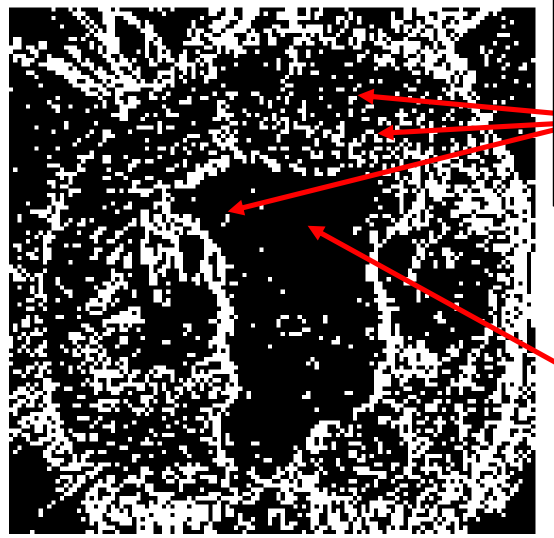
This number $\mu_{patch(i)}$is SMALLER for patch numbers 2, 3 and 4 as compared to patch numbers 1,5,6,7 and 8. There is no real discernible pattern even though the mean numbers corresponding to patches 2, 3, 4 are considerably lower as compared to the mean numbers in the remaining patches. One thing to note is that all Hampshire male pigs have white belts but female pigs need not have these belts. They do have small patches of white somewhere on their bodies.

Hampshire female with a small white patch on the frontal right leg. It does not have a white belt unlike the Hampshire male pig. This makes discrimination a lot more difficult.

Hampshire male with a prominent white belt. All Hampshire male pigs will have a white belt either partial or complete of different thickness, which could be used as a distinguishing factor for separating and tracking HAMPSHIRE MALE PIGS.

Section-2) HAIR/CILIARY FEATURES along with PORES on the MUZZLE

- What is usually not visible in the intensity profile can be extracted through the GRADIENT profiles. It is obvious that if there is sensor noise or dust present in the image this will be amplified in the gradient profiles. However these images have been captured using high resolution cameras under controlled environments (not on the field of action such as the actual dusty farmlands). Shown below is the actual image and gradient profile of Pig 41092 (Fig. 3).



Pores or locations of hair follicles on the pig.

*Note density of pores **appears** to be less on the white patch*

*Fig. 3: The image on the left is a cropped muzzle picture and the image on the right is a **gradient significance map**. Observe the number of pores/cilia in the muzzle of the pig which are not apparent in the color image on the left.*

Presence and Density of Cilia on the Muzzle

The role of the hair follicle could be multifold:

1) For ***expiration and temperature regulation***. When a cilium is present at the certain location it is indicative of a larger exposure of the inner layers of the skin;

2) ***Environmental sensing*** where the cilia could be used as ***extended antennae*** for sensing the temperature, ***wind direction***, ***extending the tactile interface*** etc;

3) Some form of ***protection***.

Stability of the CILIA/HAIR/HAIR PORES patterns over time:

It is quite possible that a majority of the hair pores and hair may remain intact for adult pigs. Reduction or increase in number of pores could be due to environmental conditions, eating and breathing habits and their own natural variations. ***The position and density of pores (coupled with their spatial bias and orientation) could be used as a biometric identifier in the case of pigs.***

Presence and Detection of CILIA/HAIR/HAIR from the white patches alone:

Observe that the gradient significance map does not show any hair pores over the white patch. In fact this is a fallacy of the detection algorithm because these pores are actually present with the same density when one zooms into the image. However since this sub-dermal extension is a specular surface the cavities are not that apparent to both the naked eye as well as the feature extraction unit.

Fig. 4: Sample muzzle images taken from different animal subjects (five variations per subject) shown above.

The feature extraction algorithm is given below:

1) Compute a Gaussian smoothened derivative along the X and Y directions by controlling the smoothing parameterσ. The discrete derivative is a discretized version of the derivative of the zero centric Gaussian function,

$$D(x) = \frac{d}{dx}\left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}\right] = -\left(\frac{x}{\sqrt{2\pi}\,\sigma^{\frac{3}{2}}}\right) e^{-\frac{x^2}{2\sigma^2}} = -c_0 x e^{-\frac{x^2}{2\sigma^2}}$$

2) Both the X and Y DISCRETE KERNELS (K_x andK_y) are generated from $D(x)$ through sub-sampling. These are $w_1 \times w_2$ matrices. In fact,

$$K_y = K_x^T$$

3) If SN is one of the snout/muzzle images as seen above and I_{SN} the intensity pattern of the snout, the intensity profile is derived from the RGB palette as follows:

$$I_{SN} = 0.3\, SN_R + 0.587\, SN_G + 0.114\, SN_B$$

Where, SN_R, SN_G and SN_B represent the RED, GREEN and BLUE channel profiles of the snout color image.

4) If $I_{SN}(i,j)$ represents the intensity level at pixel location (i,j), the X and Y gradient profiles are computed at each pixel after the image is zero padded on all sides as follows:

$$G_x(i,j) = I_{SN(pad)}(i,j) * K_x(i,j)$$
$$G_y(i,j) = I_{SN(pad)}(i,j) * K_y(i,j)$$

5) The magnitude of the smoothened gradient is computed at each pixel as:

$$M_G(i,j) = \sqrt{G_x(i,j)^2 + G_y(i,j)^2}$$

6) Compute the mean value of the gradient magnitude over the entire image:

$$\mu_G = \frac{1}{N_1 \times N_2} \sum_i \sum_j M_G(i,j)$$

where, $N_1 \times N_2$ is the size of each snout image.

7) Generate a normalized gradient profile by dividing gradient magnitude profile by the mean gradient, μ_G

$$M_{G(norm)}(i,j) = \frac{M_G(i,j)}{\mu_G}$$

8) Now construct a ***significance map*** by thresholding the normalized gradient profile. This significance map is a binary map and the threshold value is chose as UNITY. Any pixel which exhibits a normalized gradient value larger than ONE is set to ONE and the others are set to zero. This significance map is represented by a binary image of size $N_1 \times N_2$,

$$S(i,j) = \begin{cases} 1 & if\ M_{G(norm)}(i,j) > 1 \\ 0 & otherwise \end{cases}$$

The ***gradient significance maps*** for the above 25 images are shown below:

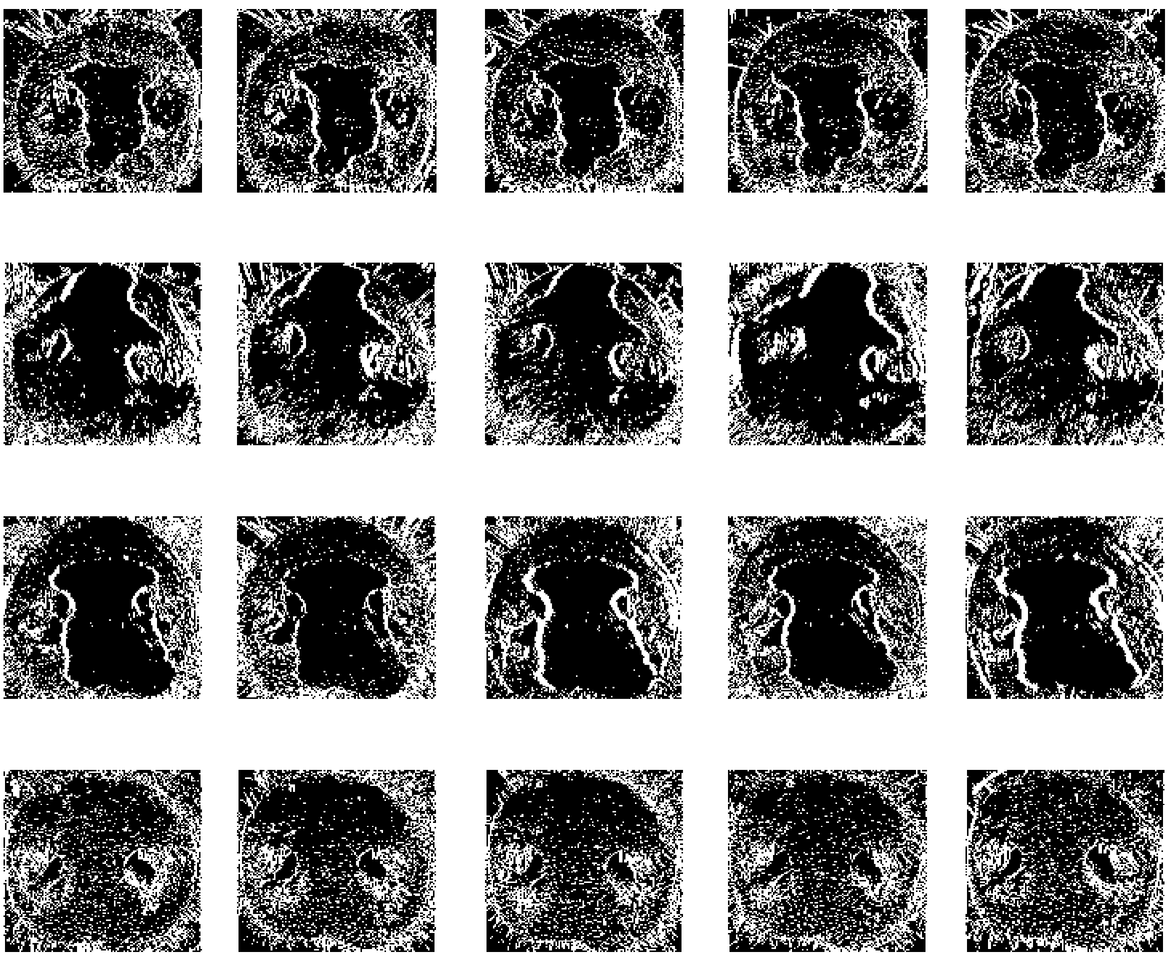


Fig. 5: Gradient significance map for all the snout images shown in Fig. 4. There is enough stability and diversity in these images to segregate the four different classes.

Taking this as the seed feature, a patch abstraction is constructed over blocks of size $N_B \times N_B$. The patch statistic is computed as shown in Fig. 6.

Patch statistics on SEED GRADIENT SIGNIFICANCE MAP

The gradient significance map is split into equal sized blocks and then a patch statistic is generated for each block. The patch statistic initially chosen is the density profile or the fraction of white dots within the block. It is obvious that this statistic is a function of gradient significance map and also the patch size N_B .

All snout images were resized to $\mathrm{N}_1 = 1000$ and $\mathrm{N}_2 = 1000$ and the patch size was varied from $N_B = 20\ to\ 250$. This secondary feature was used in the UNSUPERVISED CLUSTERING process of random snout images.

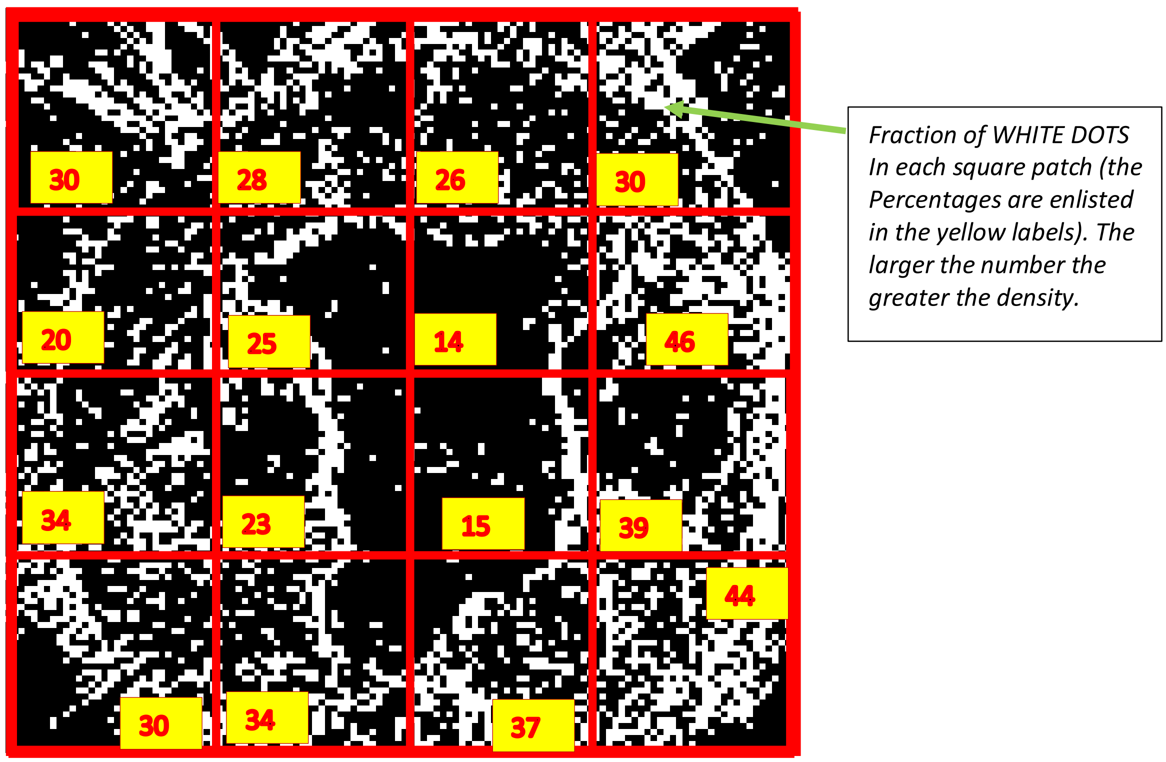


Fig. 6: Patch density profile for $N_B = 250$.

Section-3) UNSUPERVISED CLUSTERING SETUP and RESULTS

For initial testing:

- Initially only four muzzle sets were available corresponding to four distinct pigs (three probably Hampshire and one Ghungroo). The animal numbers for the four pigs were: 41092, 41093, 41097 and 41099. The number of pre-defined clusters were therefore$N_C = 4$.
- In each of these classes, $N_T = 15$ variations were selected for the analysis. The total number of images were:
$$N = N_T \times N_c = 15 \times 4 = 60$$
- Among these $N = 60$ images, $N_R = 25$ images were randomly picked for UNSUPERVISED CLUSTERING. One such random mixture is shown in Fig. 7.
- Parameter setting:
  1) The Gaussian filter was set to $\sigma = 2$ to ensure that very fine noise is masked, however the cilia and the dominant edges are clearly picked.
  2) The threshold for the normalized gradient magnitude profile as discussed earlier was set to '1'.
  3) The patch size was varied from $N_B = 20\ to\ N_B = 250; with\ N_B \in \{20, 50, 100, 200, 250\}$
  4) The k-means algorithm was used for UNSUPERVISED CLUSTERING of these N_R feature vectors from $N_R = 25$, randomly chosen snout images.

Fig. 7: Random mixture of 25 images fed to the feature extraction and UNSUPERVISED CLUSTERING UNIT.

UNSUPERVISED RESULTS FOR FEATURE VALIDATION

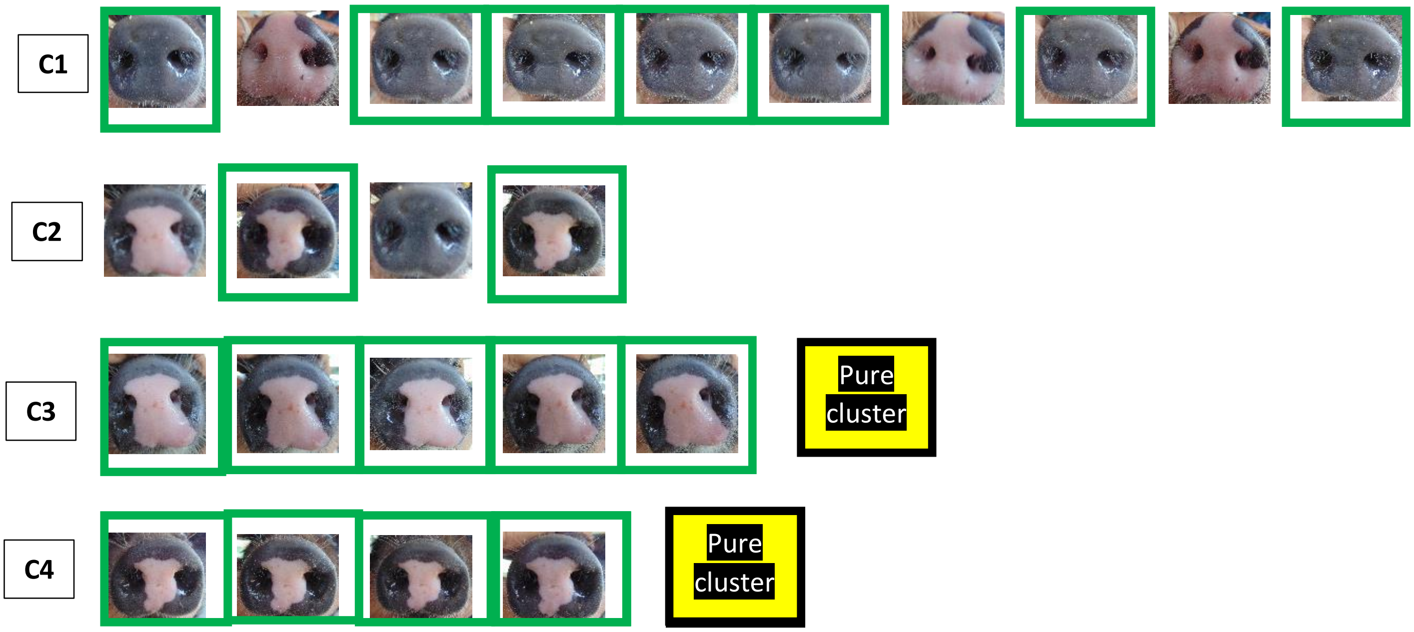


Fig. 8.1: PATCH SIZE SMALL: $N_B = 20$. Cluster purity is indicated and dominant snout patterns flagged in green rectangles. The remaining are false positives. Clusters 3 and 4 are pure with no false positives, while clusters 1 and 2 are impure with a few false positives. ***[SOME FALSE POSITIVES BUT PROMISING]***

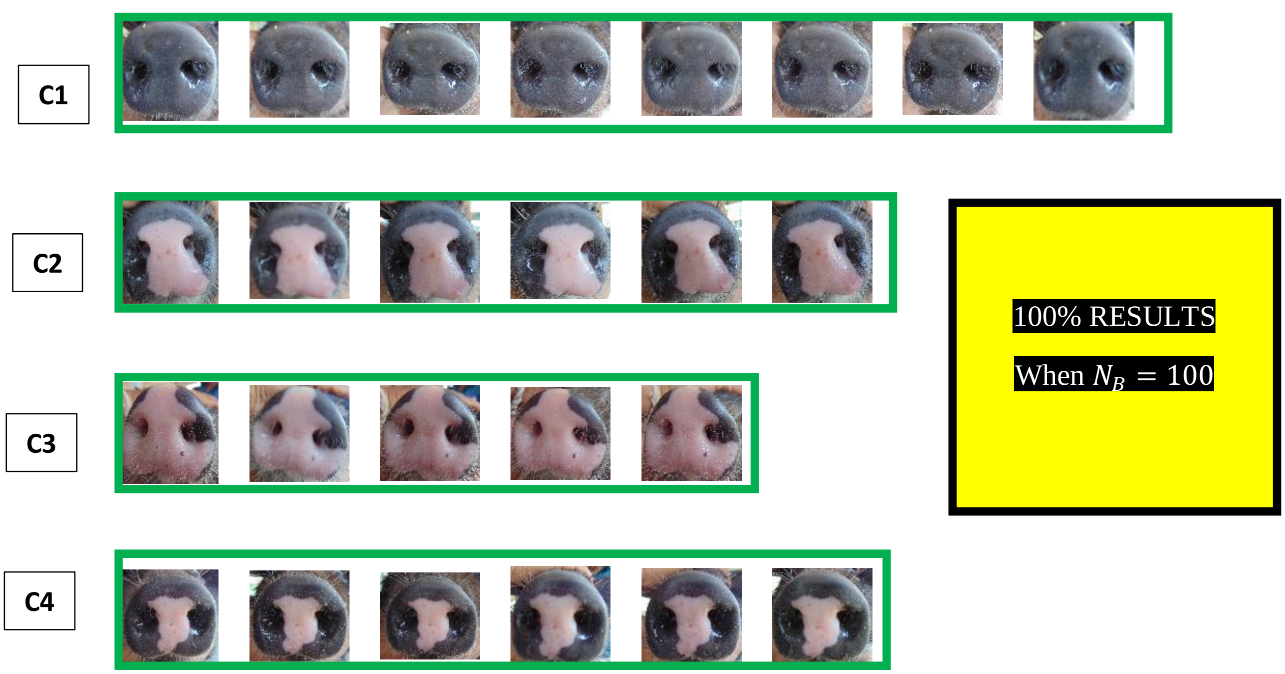


Fig. 8.2: [100% cluster purity when PATCH SIZE is increased to $N_B = 100$]

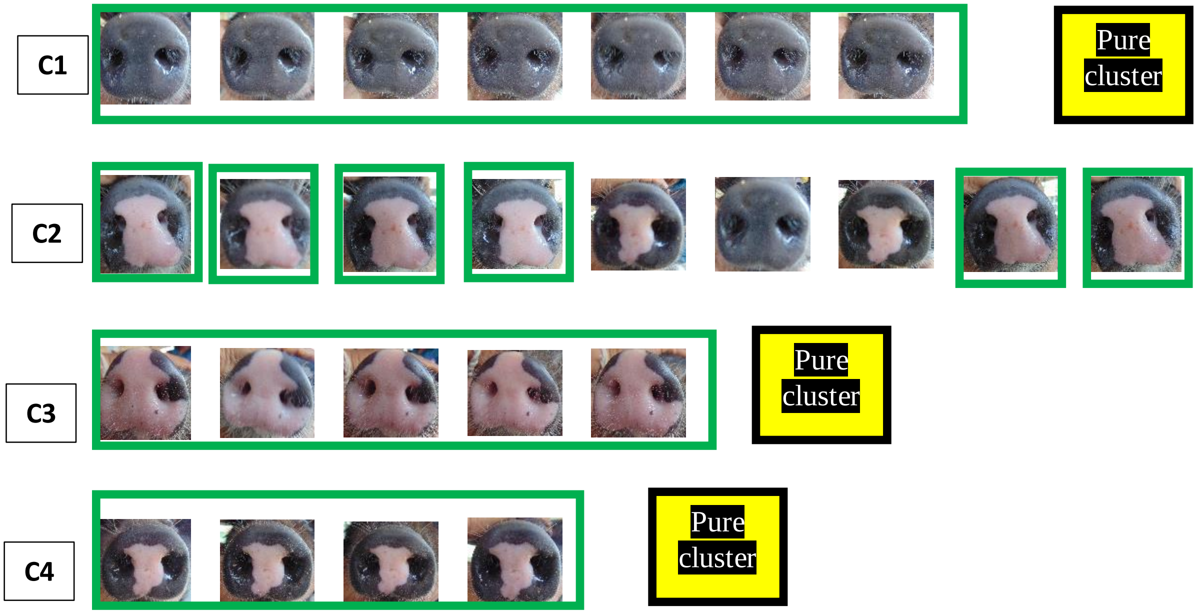


Fig. 8.3: PATCH SIZE INCREASED: $N_B = 200$. .***[SOME FALSE POSITIVES]***

Section-4) Inferences

1) The GRADIENT SIGNIFICANCE map extracted from the MUZZLE images of pigs tends to amplify the pore and ciliary arrangement on the muzzle surface, which may be expected to be distinct for each pig. This pattern can therefore be used as a BIO-IDENTIFIER for segregating individual pigs not just within the same breed but also across breeds.

2) A secondary feature was built on this gradient significance map by abstracting the details over equal size patches. This secondary feature derived from several random test images was fed to an unsupervised clustering algorithm. For certain parameter settings, classification results were a 100%. Precision drops for other settings and some clusters remain impure.

3) Skin color based analysis was inconclusive. But more rigorous testing with other morphological features such as shape contours etc., may lead to some form of convergence towards breed analysis.