Localization of Articles from a Newspaper Image

November 29, 2019

A MINI-PROJECT REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

B.Tech Degree in Computer Science and Technology

Submitted by:
Sushmita Mondal(Roll No:510518028)
Subhadeep Roy(Roll No:510518033)
Soumik Sen(Roll No:510518036)
Visesh Sharma(Roll No:510518061)
Devesh Singh(Roll No:510518073)

under the supervision of

Prof. Samit Biswas

1856 १८५६
उत्तिष्ठत जाग्रत प्राप्य वरान् निबोधत
INDIAN INSTITUTE OF ENGINEERING SCIENCE AND TECHNOLOGY, SHIBPUR
भारतीय अभियांत्रिकी विज्ञान एवं प्रौद्योगिकी संस्थान, शिवपुर

November 29, 2019

Localization of Articles from a Newspaper Image

Sushmita Biswas(Roll No:510518028)
Subhadeep Roy(Roll No:510518033)
Soumik Sen(Roll No:510518036)
Visesh Sharma(Roll No:510518061)
Devesh Singh(Roll No:510518073)

under the supervision of

Prof. Samit Biswas

Department of Computer Science and Technology

IIEST-Shibpur

India-711103

ACKNOWLEDGEMENT

We would like to express our deep gratitude to Professor Samit Biswas, Department of Computer Science and Technology,Indian Institute of Engineering Science and Technology, Shibpur, our project supervisor, for his patient guidance, enthusiastic encouragement and useful critiques for the project.

We would like to thank Prof. Sekhar Mandal, the Head of the Department, for his support.

We would also like to thank all the faculty members and staff of the Department, as well as our fellow classmates for their help and support.

(Sushmita Mondal-510518028)

(Subhadeep Roy-510518033)

(Soumik Sen-510518036)

(Visesh Sharma-510518061)

(Devesh Singh-510518073)

CERTIFICATE

This is to certify that the mini-project titled "*Segregation of Articles from a Newspaper Image*" has been carried out by **Sushmita Mondal(Roll No - 510518028), Subhadeep Roy(Roll No - 510518033), Soumik Sen(Roll No - 510518036), Visesh Sharma(Roll No - 510518061) and Devesh Singh(Roll No - 510518073)** under my supervision and guidance in their third and fourth semesters in partial fulfillment of the requirements for the B.Tech Programme in the Department of Computer Science and Technology, IIEST Shibpur.This work has not been reported anywhere for any other purpose.

(Samit Biswas)

Contents

List of Figures

A INTRODUCTION

The newspaper is the most powerful of all the means of expression of the news and views about men and matters.Newspapers are regarded by economists as a necessity of modern life. With the growth of literacy and the development of the means of communication they are playing a very important role in nowadays' society.

The functions performed by newspapers cannot be over-emphasized in any society in general and in academic institution in particular. So several organizations like libraries and colleges try to preserve the important articles of old newspapers. But they encounter several problems in managing newspapers.

Newspaper preservation is a challenge because newsprint is an inherently unstable paper. Formulated to be inexpensive and expendable, newsprint is manufactured with large percentages of unpurified wood pulp which contains impurities that remain in the paper after processing.These impurities when exposed to light, high humidity and atmospheric pollutants, promote discoloration and acidic reactions in the paper. Acidity causes the paper fibres to weaken and break, and is the major culprit in causing the paper to become brittle.All these shortcomings make newspaper preservation a tough challenge.

When newspapers or clippings are valued most for the information they contain, and not as artifacts, copying the information onto a more permanent quality paper should by undertaken.Still they can decay. But an image of a newspaper will remain as long as it is kept. So taking images is a far better alternative of preserving old newspapers.

The project mainly focuses on separating the different articles present in a newspaper image so that they can be analysed and segregated under different genres.

A.1 Motivation

The budget that goes for the management of newspapers in an average library is a lot yet it is not enough to properly preserve the newspapers. Libraries encounter several problems in managing newspapers. Librarians are confronted with the problems of handling the papers because of their fragility; preserving them because of their information contents, and housing them because of lack of space and the rapid rate at which they are churned out. The system of storage and retrieval increases the incidence of destroying and or tearing the newspapers most especially when tying and untying the papers.

Also these newspapers are needed by readers who in times feel the nessecity to take them home. But this causes the newspapers to be damaged making it more and more hard to preserve. Even searching old articles become tough and time consuming for the fragility of old newspapers.

Therefore, the motivation behind the project is to simplify the process of newspaper preservation and make it more economical.

A.2 The Proposed Project

The programme that is being proposed under this project takes the image of a newspaper page and separates all the articles present in it. It then makes images of all these articles and saves them.

A.3 Outline Of Project

Section 2: provides the details of all the programmes and applications used to make the project.

Section 3: covers the programme written and all the methods written.

Section 4: discusses the limitations of the project and how it can be improved.

Section 5: discusses how the project can be used.

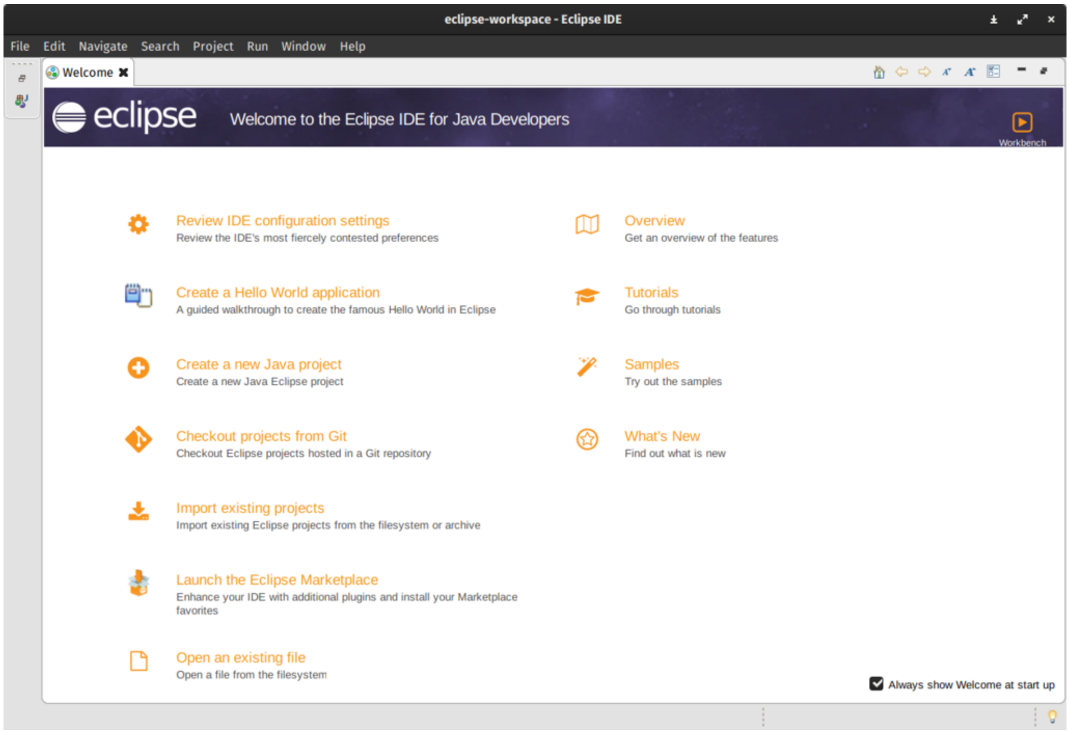
Section 6: References.

B PREREQUISITES

All the programmes and the applications used to make the project are listed below:

B.1 Eclipse IDE

Eclipse is an Integrated Development Environment (IDE) used in computer programming. It contains a base workspace and an extensible plug-in system for customizing the environment. Since the final project is written in JAVA Programming Language, Eclipse is used as the IDE.


eclipse-workspace - Eclipse IDE
File Edit Navigate Search Project Run Window Help
Welcome
eclipse
Workbench

Welcome to the Eclipse IDE for Java Developers

Review IDE configuration settings
Review the IDE's most fiercely contested preferences

Create a Hello World application
A guided walkthrough to create the famous Hello World in Eclipse

Create a new Java project
Create a new Java Eclipse project

Checkout projects from Git
Checkout Eclipse projects hosted in a Git repository

Import existing projects
Import existing Eclipse projects from the filesystem or archive

Launch the Eclipse Marketplace
Enhance your IDE with additional plugins and install your Marketplace favorites

Open an existing file
Open a file from the filesystem

Overview
Get an overview of the features

Tutorials
Go through tutorials

Samples
Try out the samples

What's New
Find out what is new

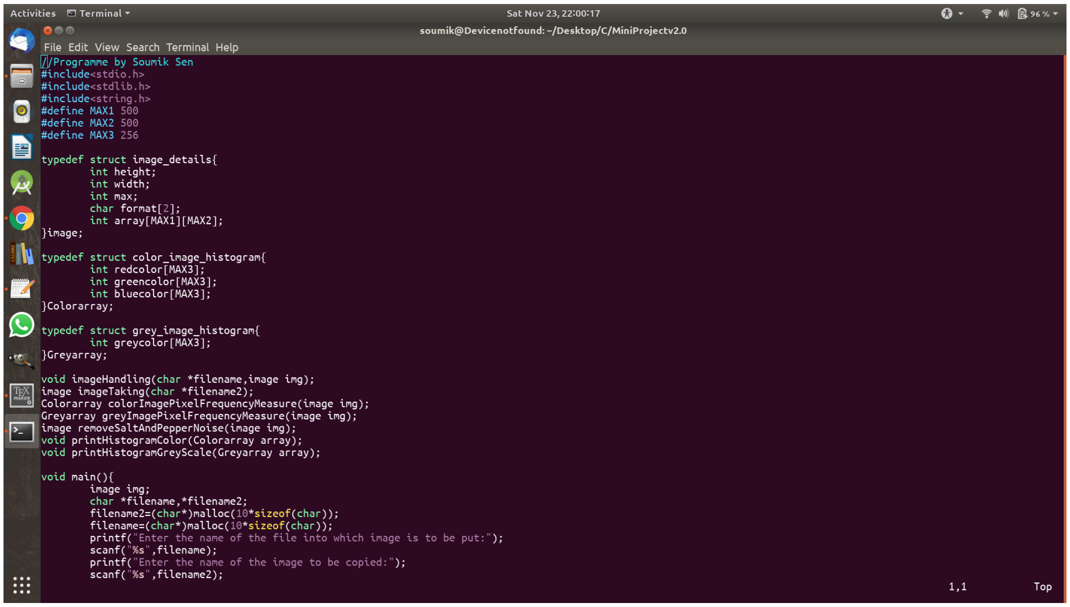
☑ Always show Welcome at start up

Figure 1: Eclipse IDE

B.2 Knowledge about images and Image Handling

To get a little amount of knowledge about images and image handling programmetically a side project was done. This project consisted of writing a programme in C using VI Text Editor. The programme takes an .netpbm format image as input does the following operations:

1. **Duplicate an Image**
2. **Making a Histogram of all the colours used in the Image**
3. **Removing Salt and Pepper Noise from the Image**


Activities
Terminal
Sat Nov 23, 22:00:17
96 %
soumik@Devicenotfound: ~/Desktop/C/MiniProjectv2.0
File Edit View Search Terminal Help

```c
//Programme by Soumik Sen
#include<stdio.h>
#include<stdlib.h>
#include<string.h>
#define MAX1 500
#define MAX2 500
#define MAX3 256

typedef struct image_details{
        int height;
        int width;
        int max;
        char format[2];
        int array[MAX1][MAX2];
}image;

typedef struct color_image_histogram{
        int redcolor[MAX3];
        int greencolor[MAX3];
        int bluecolor[MAX3];
}Colorarray;

typedef struct grey_image_histogram{
        int greycolor[MAX3];
}Greyarray;

void imageHandling(char *filename,image img);
image imageTaking(char *filename2);
Colorarray colorImagePixelFrequencyMeasure(image img);
Greyarray greyImagePixelFrequencyMeasure(image img);
image removeSaltAndPepperNoise(image img);
void printHistogramColor(Colorarray array);
void printHistogramGreyScale(Greyarray array);

void main(){
        image img;
        char *filename,*filename2;
        filename2=(char*)malloc(10*sizeof(char));
        filename=(char*)malloc(10*sizeof(char));
        printf("Enter the name of the file into which image is to be put:");
        scanf("%s",filename);
        printf("Enter the name of the image to be copied:");
        scanf("%s",filename2);
                                                                                          1,1           Top
```

Figure 2: Code picture

B.3 LaTeX

LaTeX is a document preparation system. When writing, the writer uses plain text as opposed to the formatted text found in WYSIWYG word processors like Microsoft Word, LibreOffice Writer and Apple Pages. The writer uses markup tagging conventions to define the general structure of a document. LaTeX with TEXMAKER as IDE has been used to write the project document.

Figure 3: Latex Logo

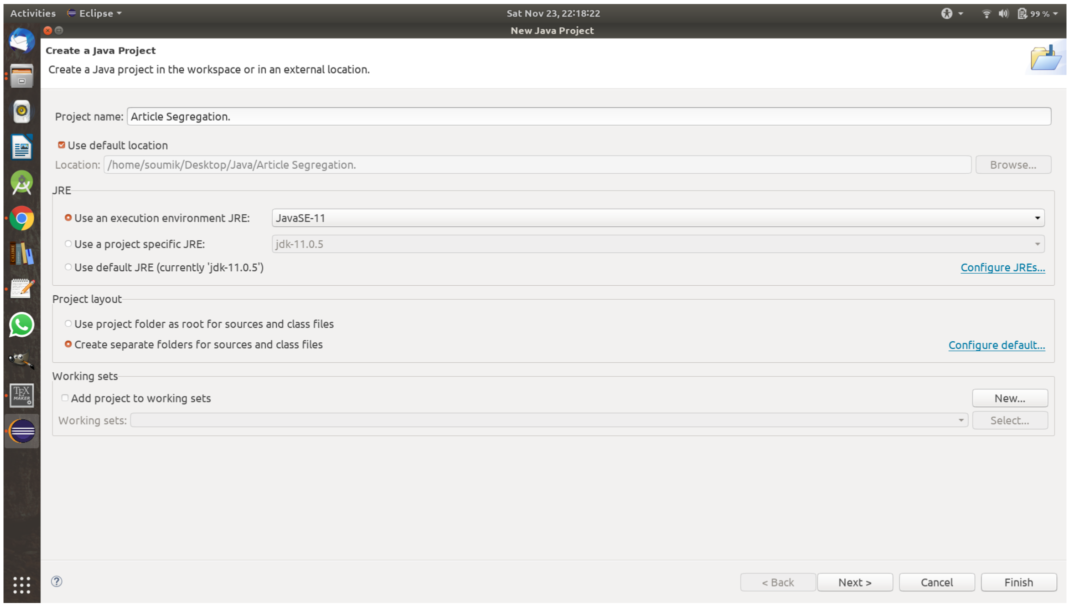
C PROGRAMME

The programme has been written in JAVA 11 Programming Language on Eclipse IDE. The following steps were followed:

C.1 Creating Project

Step 1: In the *Welcome to the Eclipse IDE for Java Developers* window, Click *Create a new Java Project.*

Step 2: In the New Java Project screen, enter the following values:
Project name: Article Segregation
Use an execution environment JRE: JavaSE-11



New Java Project

Create a Java Project

Create a Java project in the workspace or in an external location.

Project name: Article Segregation.

☑ Use default location

Location: /home/soumik/Desktop/Java/Article Segregation.

JRE

● Use an execution environment JRE: JavaSE-11

○ Use a project specific JRE: jdk-11.0.5

○ Use default JRE (currently 'jdk-11.0.5')

Project layout

○ Use project folder as root for sources and class files

● Create separate folders for sources and class files

Working sets

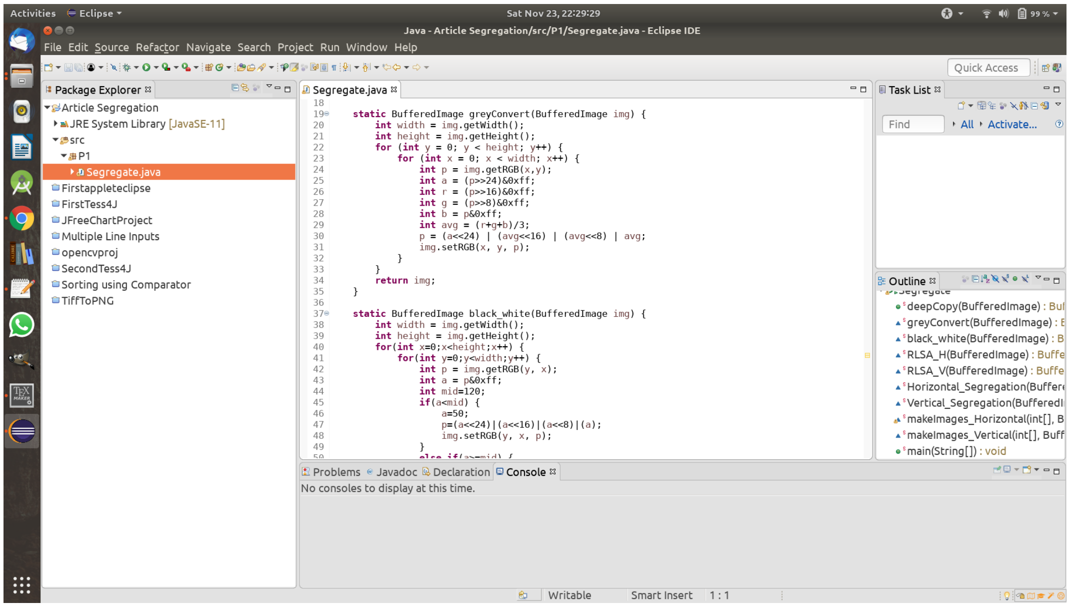
□ Add project to working sets

Working sets:

Figure 4: Creating Project

Finally click finish project.

C.2 Article Segmentation



Activities
Eclipse
Sat Nov 23, 22:29:29
99 %
Java - Article Segregation/src/P1/Segregate.java - Eclipse IDE
File Edit Source Refactor Navigate Search Project Run Window Help
Quick Access
Package Explorer
Article Segregation
JRE System Library [JavaSE-11]
src
P1
Segregate.java
Firstappleteclipse
FirstTess4J
JFreeChartProject
Multiple Line Inputs
opencvproj
SecondTess4J
Sorting using Comparator
TiffToPNG
Segregate.java
Task List
Find
All
Activate...
Outline
Segregate
deepCopy(BufferedImage) : Buf
greyConvert(BufferedImage) : B
black_white(BufferedImage) : B
RLSA_H(BufferedImage) : Buffe
RLSA_V(BufferedImage) : Buffe
Horizontal_Segregation(Buffere
Vertical_Segregation(BufferedI
makeImages_Horizontal(int[], B
makeImages_Vertical(int[], Buff
main(String[]) : void
Problems
Javadoc
Declaration
Console
Writable
Smart Insert
1 : 1

```java

    static BufferedImage greyConvert(BufferedImage img) {
        int width = img.getWidth();
        int height = img.getHeight();
        for (int y = 0; y < height; y++) {
            for (int x = 0; x < width; x++) {
                int p = img.getRGB(x,y);
                int a = (p>>24)&0xff;
                int r = (p>>16)&0xff;
                int g = (p>>8)&0xff;
                int b = p&0xff;
                int avg = (r+g+b)/3;
                p = (a<<24) | (avg<<16) | (avg<<8) | avg;
                img.setRGB(x, y, p);
            }
        }
        return img;
    }

    static BufferedImage black_white(BufferedImage img) {
        int width = img.getWidth();
        int height = img.getHeight();
        for(int x=0;x<height;x++) {
            for(int y=0;y<width;y++) {
                int p = img.getRGB(y, x);
                int a = p&0xff;
                int mid=120;
                if(a<mid) {
                    a=50;
                    p=(a<<24)|(a<<16)|(a<<8)|(a);
                    img.setRGB(y, x, p);
                }
                else if(a>=mid) {
```

No consoles to display at this time.

Figure 5: Project Code

The left tab shows the Project Structure. The right tab named *Outline* shows the different methods used in the project.

C.3 Explaining the Methods

deepCopy: This method copies the contents of a BufferedImage object into another BufferedImage object.

black-white: This method converts a colour image into a black and white image.

Horizontal-Segregation: This method cuts the image into horizontal segments according to white spaces.

Vertical-Segregation: This method cuts the image into vertical segments according to white spaces.

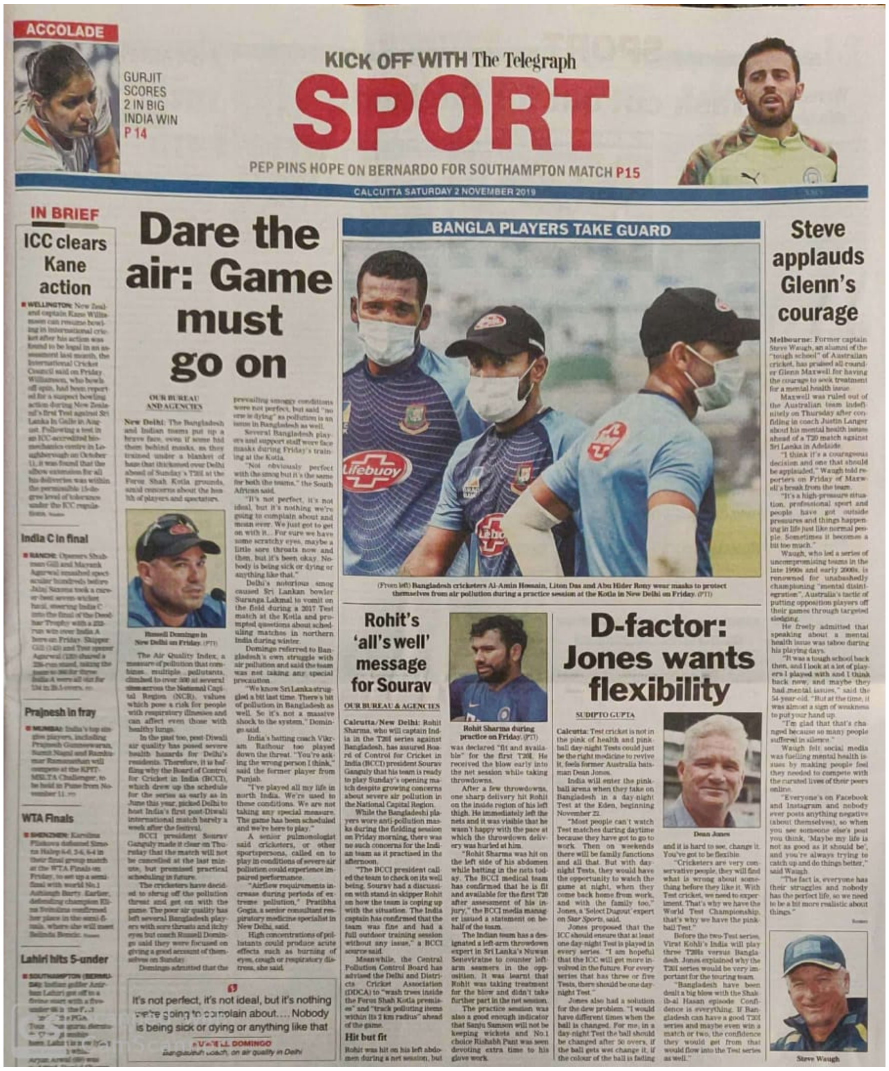
makeImages-Vertical: This creates small subimages of the vertically segmented images obtained from *Vertical-Segregation* method.

makeImages-Horizontal: This creates small subimages of the horizontally segmented images obtained from *Horizontal-Segregation* method.

C.4 Process of Separation

The main idea of the project is to separate the different articles of the input images. To do so we try to identify the articles by the blank spaces between them. If there is a horizontal space such that it reaches from left end to right end then the *Horizontal-Segregation* method cuts the image there. Similarly if there is a vertical space that extends from the top to the bottom of the image then the *Vertical-Segregation* method cuts the image. Then all the cut pieces are converted into images by the *makeImage* methods.

Example:



ACCOLADE

GURJIT SCORES 2 IN BIG INDIA WIN P 14

KICK OFF WITH The Telegraph

SPORT

PEP PINS HOPE ON BERNARDO FOR SOUTHAMPTON MATCH P15

CALCUTTA SATURDAY 2 NOVEMBER 2019

IN BRIEF

ICC clears Kane action

WELLINGTON: New Zealand captain Kane Williamson can resume bowling in international cricket after his action was found to be legal in an assessment last month, the International Cricket Council said on Friday. Williamson, who bowls off spin, had been reported for a suspect bowling action during New Zealand's first Test against Sri Lanka in Galle in August. Following a test in an ICC-accredited biomechanics centre in Loughborough on October 11, it was found that the elbow extension for all his deliveries was within the permissible 15-degree level of tolerance under the ICC regulations. Reuters

India C in final

RANCHI: Openers Shubman Gill and Mayank Agarwal smashed spectacular hundreds before Jalaj Saxena took a career-best seven-wicket haul, steering India C into the final of the Deodhar Trophy with a 232-run win over India A here on Friday. Skipper Gill (143) and Test opener Agarwal (120) shared a 226-run stand, taking the team to 366 for three. India A were all out for 134 in 29.5 overs. PTI

Prajnesh in fray

MUMBAI: India's top singles players, including Prajnesh Gunneswaran, Sumit Nagal and Ramkumar Ramanathan will compete at the KPIT-MSLTA Challenger, to be held in Pune from November 11. PTI

WTA Finals

SHENZHEN: Karolina Pliskova defeated Simona Halep 6-0, 2-6, 6-4 in their final group match at the WTA Finals on Friday, to set up a semi-final with world No.1 Ashleigh Barty. Earlier, defending champion Elina Svitolina confirmed her place in the semi-finals, where she will meet Belinda Bencic. Reuters

Lahiri hits 5-under

SOUTHAMPTON (BERMUDA): Indian golfer Anirban Lahiri got off to a flying start with a five-under 66 in the [illegible] PGA Tour [illegible] Aryan Arwal (69) was [illegible]

Dare the air: Game must go on

OUR BUREAU AND AGENCIES

New Delhi: The Bangladesh and Indian teams put up a brave face, even if some hid them behind masks, as they trained under a blanket of haze that thickened over Delhi ahead of Sunday's T20I at the Feroz Shah Kotla grounds, amid concerns about the health of players and spectators.

Russell Domingo in New Delhi on Friday. (PTI)

The Air Quality Index, a measure of pollution that combines multiple pollutants, climbed to over 500 at several sites across the National Capital Region (NCR), values which pose a risk for people with respiratory illnesses and can affect even those with healthy lungs.

In the past too, post-Diwali air quality has posed severe health hazards for Delhi's residents. Therefore, it is baffling why the Board of Control for Cricket in India (BCCI), which drew up the schedule for the series as early as in June this year, picked Delhi to host India's first post-Diwali international match barely a week after the festival.

BCCI president Sourav Ganguly made it clear on Thursday that the match will not be cancelled at the last minute, but promised practical scheduling in future.

The cricketers have decided to shrug off the pollution threat and get on with the game. The poor air quality has left several Bangladesh players with sore throats and itchy eyes but coach Russell Domingo said they were focused on giving a good account of themselves on Sunday.

Domingo admitted that the prevailing smoggy conditions were not perfect, but said "no one is dying" as pollution is an issue in Bangladesh as well.

Several Bangladesh players and support staff wore face masks during Friday's training at the Kotla.

"Not obviously perfect with the smog but it's the same for both the teams," the South African said.

"It's not perfect, it's not ideal, but it's nothing we're going to complain about and moan ever. We just got to get on with it... For sure we have some scratchy eyes, maybe a little sore throats now and then, but it's been okay. Nobody is being sick or dying or anything like that."

Delhi's notorious smog caused Sri Lankan bowler Suranga Lakmal to vomit on the field during a 2017 Test match at the Kotla and prompted questions about scheduling matches in northern India during winter.

Domingo referred to Bangladesh's own struggle with air pollution and said the team was not taking any special precaution.

"We know Sri Lanka struggled a bit last time. There's bit of pollution in Bangladesh as well. So it's not a massive shock to the system," Domingo said.

India's batting coach Vikram Rathour too played down the threat. "You're asking the wrong person I think," said the former player from Punjab.

"I've played all my life in north India. We're used to these conditions. We are not taking any special measure. The game has been scheduled and we're here to play."

A senior pulmonologist said cricketers, or other sportspersons, called on to play in conditions of severe air pollution could experience impaired performance.

"Airflow requirements increase during periods of extreme pollution," Pratibha Gogia, a senior consultant respiratory medicine specialist in New Delhi, said.

High concentrations of pollutants could produce acute effects such as burning of eyes, cough or respiratory distress, she said.

It's not perfect, it's not ideal, but it's nothing we're going to complain about.... Nobody is being sick or dying or anything like that

RUSSELL DOMINGO

Bangladesh coach, on air quality in Delhi

BANGLA PLAYERS TAKE GUARD

(From left) Bangladesh cricketers Al-Amin Hossain, Liton Das and Abu Hider Rony wear masks to protect themselves from air pollution during a practice session at the Kotla in New Delhi on Friday. (PTI)

Rohit's 'all's well' message for Sourav

OUR BUREAU & AGENCIES

Calcutta/New Delhi: Rohit Sharma, who will captain India in the T20I series against Bangladesh, has assured Board of Control for Cricket in India (BCCI) president Sourav Ganguly that his team is ready to play Sunday's opening match despite growing concerns about severe air pollution in the National Capital Region.

While the Bangladeshi players wore anti-pollution masks during the fielding session on Friday morning, there was no such concerns for the Indian team as it practised in the afternoon.

"The BCCI president called the team to check on its well being. Sourav had a discussion with stand-in skipper Rohit on how the team is coping up with the situation. The India captain has confirmed that the team was fine and had a full outdoor training session without any issue," a BCCI source said.

Meanwhile, the Central Pollution Control Board has advised the Delhi and Districts Cricket Association (DDCA) to "wash trees inside the Feroz Shah Kotla premises" and "track polluting items within its 2 km radius" ahead of the game.

Hit but fit

Rohit Sharma during practice on Friday. (PTI)

Rohit was hit on his left abdomen during a net session, but was declared "fit and available" for the first T20I. He received the blow early into the net session while taking throwdowns.

After a few throwdowns, one sharp delivery hit Rohit on the inside region of his left thigh. He immediately left the nets and it was visible that he wasn't happy with the pace at which the throwdown delivery was hurled at him.

"Rohit Sharma was hit on the left side of his abdomen while batting in the nets today. The BCCI medical team has confirmed that he is fit and available for the first T20 after assessment of his injury," the BCCI media manager issued a statement on behalf of the team.

The Indian team has a designated a left-arm throwdown expert in Sri Lanka's Nuwan Seneviratne to counter left-arm seamers in the opposition. It was learnt that Rohit was taking treatment for the blow and didn't take further part in the net session.

The practice session was also a good enough indicator that Sanju Samson will not be keeping wickets and No.1 choice Rishabh Pant was seen devoting extra time to his glove work.

D-factor: Jones wants flexibility

SUDIPTO GUPTA

Calcutta: Test cricket is not in the pink of health and pink-ball day-night Tests could just be the right medicine to revive it, feels former Australia batsman Dean Jones.

India will enter the pink-ball arena when they take on Bangladesh in a day-night Test at the Eden, beginning November 22.

"Most people can't watch Test matches during daytime because they have got to go to work. Then on weekends there will be family functions and all that. But with day-night Tests, they would have the opportunity to watch the game at night, when they come back home from work, and with the family too," Jones, a 'Select Dugout' expert on *Star Sports*, said.

Jones proposed that the ICC should ensure that at least one day-night Test is played in every series. "I am hopeful that the ICC will get more involved in the future. For every series that has three or five Tests, there should be one day-night Test."

Jones also had a solution for the dew problem. "I would have different times when the ball is changed. For me, in a day-night Test the ball should be changed after 50 overs, if the ball gets wet change it, if the colour of the ball is fading and it is hard to see, change it. You've got to be flexible.

Dean Jones

"Cricketers are very conservative people, they will find what is wrong about something before they like it. With Test cricket, we need to experiment. That's why we have the World Test Championship, that's why we have the pink-ball Test."

Before the two-Test series, Virat Kohli's India will play three T20Is versus Bangladesh. Jones explained why the T20I series would be very important for the touring team.

"Bangladesh have been dealt a big blow with the Shakib-al Hasan episode. Confidence is everything. If Bangladesh can have a good T20I series and maybe even win a match or two, the confidence they would get from that would flow into the Test series as well."

Steve applauds Glenn's courage

Melbourne: Former captain Steve Waugh, an alumni of the "tough school" of Australian cricket, has praised all-rounder Glenn Maxwell for having the courage to seek treatment for a mental health issue.

Maxwell was ruled out of the Australian team indefinitely on Thursday after confiding in coach Justin Langer about his mental health issues ahead of a T20 match against Sri Lanka in Adelaide.

"I think it's a courageous decision and one that should be applauded," Waugh told reporters on Friday of Maxwell's break from the team.

"It's a high-pressure situation, professional sport and people have got outside pressures and things happening in life just like normal people. Sometimes it becomes a bit too much."

Waugh, who led a series of uncompromising teams in the late 1990s and early 2000s, is renowned for unabashedly championing "mental disintegration", Australia's tactic of putting opposition players off their games through targeted sledging.

He freely admitted that speaking about a mental health issue was taboo during his playing days.

"It was a tough school back then, and I look at a lot of players I played with and I think back now, and maybe they had mental issues," said the 54-year-old. "But at the time, it was almost a sign of weakness to put your hand up.

"I'm glad that that's changed because so many people suffered in silence."

Waugh felt social media was fuelling mental health issues by making people feel they needed to compete with the curated lives of their peers online.

"Everyone's on Facebook and Instagram and nobody ever posts anything negative (about themselves), so when you see someone else's post you think, 'Maybe my life is not as good as it should be', and you're always trying to catch up and do things better," said Waugh.

"The fact is, everyone has their struggles and nobody has the perfect life, so we need to be a bit more realistic about things."

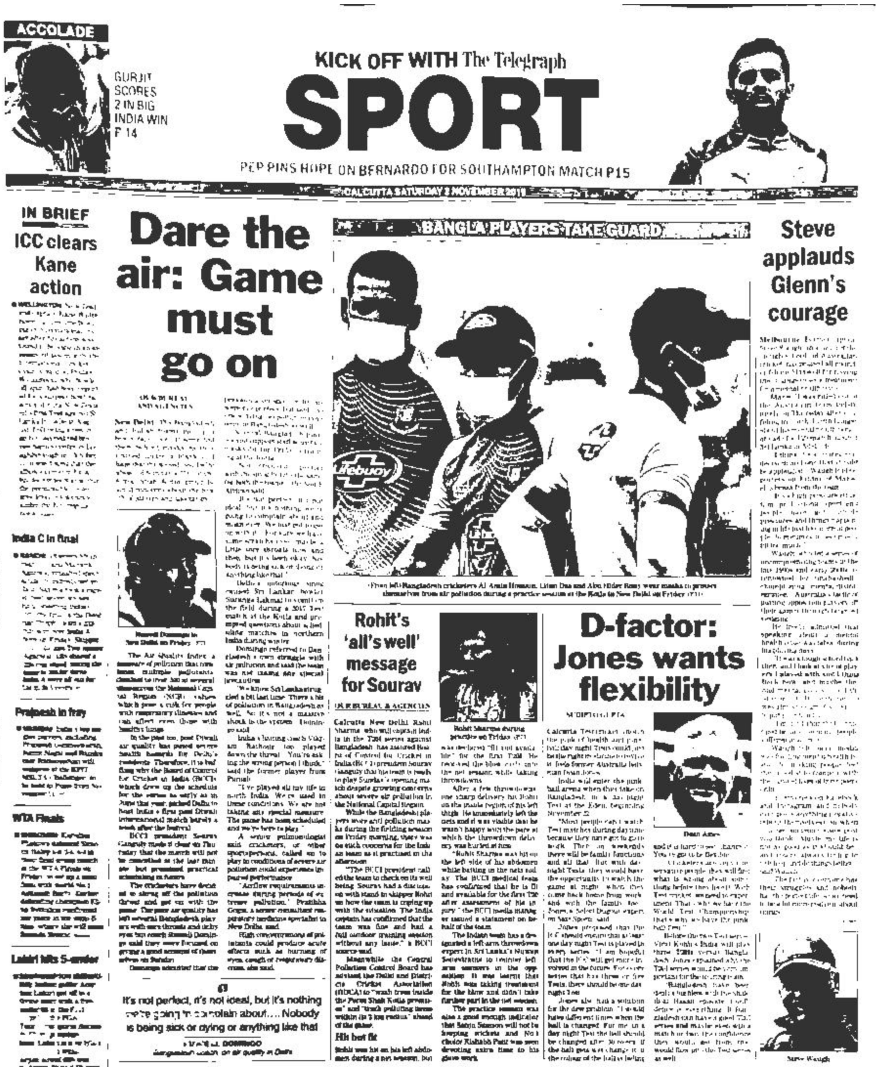
Reuters

Steve Waugh

Figure 6: Input Image

We first convert the colour image to binary image by taking 120 pixel as a midvalue and making any value above 120 pixel to 255 pixel(white) and all values below 120 pixel to 0 pixel(black).

We take 120 as midvalue as it is the median of all the pixels present in the image.



ACCOLADE

GURJIT SCORES 2 IN BIG INDIA WIN P 14

KICK OFF WITH The Telegraph

SPORT

PEP PINS HOPE ON BERNARDO FOR SOUTHAMPTON MATCH P15

CALCUTTA SATURDAY 2 NOVEMBER 2019

IN BRIEF

ICC clears Kane action

India C in final

Prajnesh in fray

WTA Finals

Lahiri hits 5-under

Dare the air: Game must go on

Russell Domingo in New Delhi on Friday

It's not perfect, it's not ideal, but it's nothing we're going to complain about.... Nobody is being sick or dying or anything like that

RUSSELL DOMINGO
Bangladesh coach on air quality in Delhi

BANGLA PLAYERS TAKE GUARD

(From left) Bangladesh cricketers Al Amin Hossain, Liton Das and Abu Hider Rony wear masks to protect themselves from air pollution during a practice session at the Kotla in New Delhi on Friday

Rohit's 'all's well' message for Sourav

OUR BUREAU & AGENCIES

Rohit Sharma during practice on Friday

His hot fix

D-factor: Jones wants flexibility

Dean Jones

Steve applauds Glenn's courage

Steve Waugh

Figure 7: Binary Image Conversion

After the Conversion to Black and White Image we first cut the image horizontally. To show that we use the *horizontal summation graph* of the image.

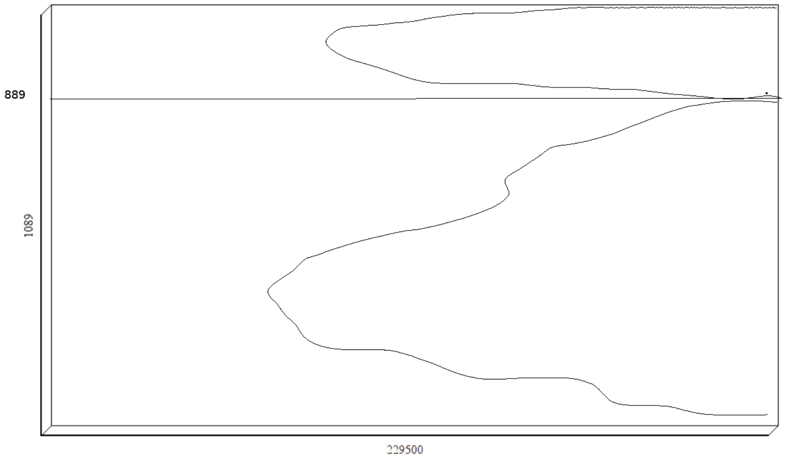


Figure 8: Horizontal Graph

This graph has 900*255 = 229500 as the x-axis length. This value is the value that comes when we add all the white pixels in a row of length 900. So the image has 3 such rows at 0 height, at 889 height and at 1089 height from the bottom. The segmented images are :

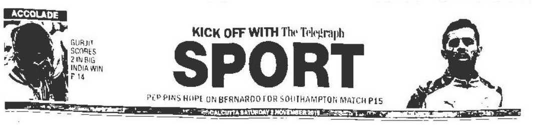
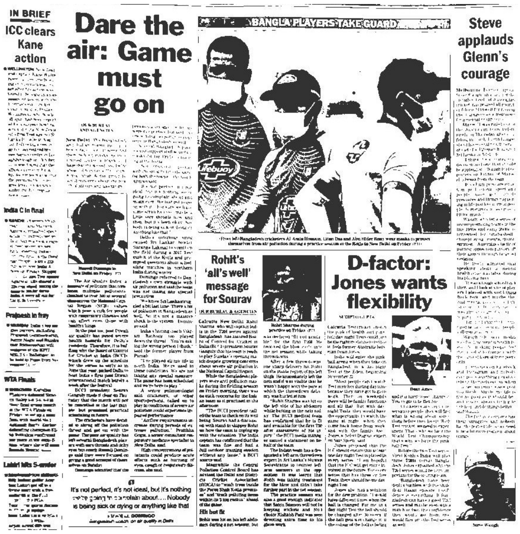

ACCOLADE
GURJIT SCORES 2 IN BIG INDIA WIN P 14
KICK OFF WITH The Telegraph
SPORT
PEP PINS HOPE ON BERNARDO FOR SOUTHAMPTON MATCH P15

Figure 9: First Segment



IN BRIEF

ICC clears Kane action

India C in final

Prajnesh in fray

WTA Finals

Lahiri hits 5-under

Dare the air: Game must go on

It's not perfect, it's not ideal, but it's nothing we're going to complain about.... Nobody is being sick or dying or anything like that

BANGLA PLAYERS TAKE GUARD

Steve applauds Glenn's courage

Rohit's 'all's well' message for Sourav

D-factor: Jones wants flexibility

Figure 10: Second Segment

Now we localize the two segments vertically. To show that we use the *vertical summation graph* of the image.

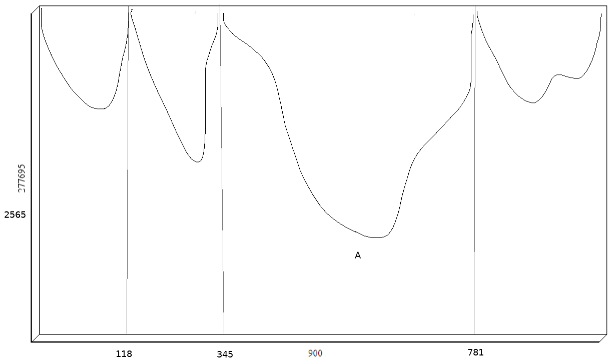
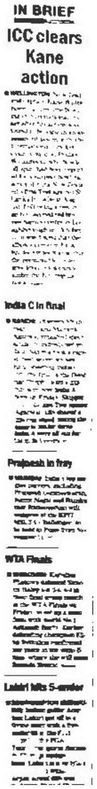


Figure 11: Vertical graph

The y axis has maximum value of 1089*255 = 277695 pixels which is the summation value when an entire column of the image is white. So whenever the value reaches this maximum value there is a white column and the image is segmented from there, similar to horizontal segmentation. The resulted segments are given below:



IN BRIEF

ICC clears Kane action

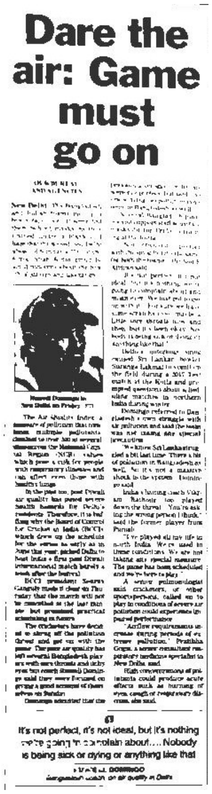
India C in final

Prajnesh in fray

WTA Finals

Lahiri hits 5-under

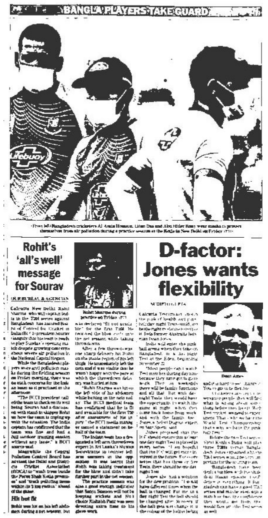
(a) Segment



Dare the air: Game must go on

It's not perfect, it's not ideal, but it's nothing we're going to complain about.... Nobody is being sick or dying or anything like that

(b) Segment

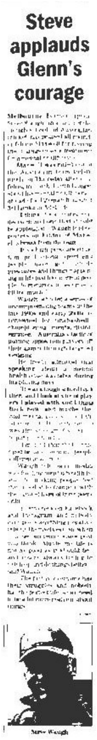


BANGLA PLAYERS TAKE GUARD

Rohit's 'all's well' message for Sourav

D-factor: Jones wants flexibility

(c) Segment



Steve applauds Glenn's courage

(d) Segment

Finally we collect all the segments and recreate them from the original colour image to get all the localized articles. Thus, the algorithm works in this manner.

RESULTS:

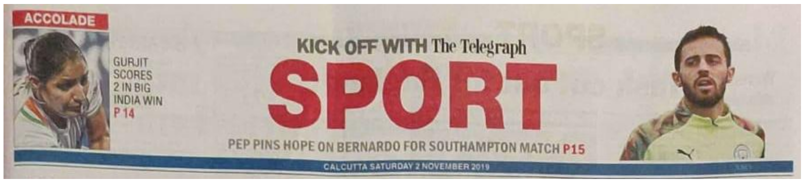
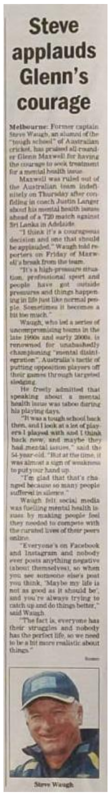

ACCOLADE
GURJIT SCORES 2 IN BIG INDIA WIN P 14
KICK OFF WITH The Telegraph
SPORT
PEP PINS HOPE ON BERNARDO FOR SOUTHAMPTON MATCH P15
CALCUTTA SATURDAY 2 NOVEMBER 2019

Figure 13: The Section



Steve applauds Glenn's courage

Melbourne: Former captain Steve Waugh, an alumni of the "tough school" of Australian cricket, has praised all-rounder Glenn Maxwell for having the courage to seek treatment for a mental health issue.

Maxwell was ruled out of the Australian team indefinitely on Thursday after confiding in coach Justin Langer about his mental health issues ahead of a T20 match against Sri Lanka in Adelaide.

"I think it's a courageous decision and one that should be applauded," Waugh told reporters on Friday of Maxwell's break from the team.

"It's a high-pressure situation, professional sport and people have got outside pressures and things happening in life just like normal people. Sometimes it becomes a bit too much."

Waugh, who led a series of uncompromising teams in the late 1990s and early 2000s, is renowned for unabashedly championing "mental disintegration", Australia's tactic of putting opposition players off their games through targeted sledging.

He freely admitted that speaking about a mental health issue was taboo during his playing days.

"It was a tough school back then, and I look at a lot of players I played with and I think back now, and maybe they had mental issues," said the 54-year-old. "But at the time, it was almost a sign of weakness to put your hand up.

"I'm glad that that's changed because so many people suffered in silence."

Waugh felt social media was fuelling mental health issues by making people feel they needed to compete with the curated lives of their peers online.

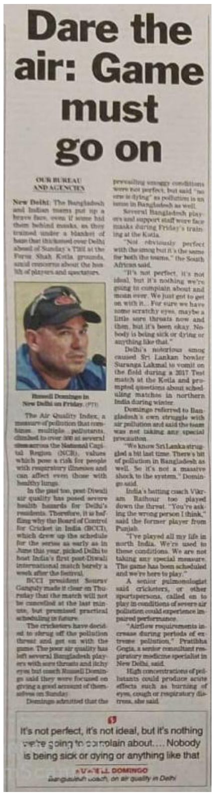
"Everyone's on Facebook and Instagram and nobody ever posts anything negative (about themselves), so when you see someone else's post you think, 'Maybe my life is not as good as it should be', and you're always trying to catch up and do things better," said Waugh.

"The fact is, everyone has their struggles and nobody has the perfect life, so we need to be a bit more realistic about things."

Reuters

Steve Waugh

(a) Segment



Dare the air: Game must go on

OUR BUREAU AND AGENCIES

New Delhi: The Bangladesh and Indian teams put up a brave face, even if some had them behind masks, as they trained under a blanket of haze that thickened over Delhi ahead of Sunday's T20I at the Feroz Shah Kotla grounds, amid concerns about the health of players and spectators.

Russell Domingo in New Delhi on Friday. (PTI)

The Air Quality Index, a measure of pollution that combines multiple pollutants, climbed to over 300 at several sites across the National Capital Region (NCR), values which pose a risk for people with respiratory illnesses and can affect even those with healthy lungs.

In the past too, post-Diwali air quality has posed severe health hazards for Delhi's residents. Therefore, it is baffling why the Board of Control for Cricket in India (BCCI), which drew up the schedule for the series as early as in June this year, picked Delhi to host India's first post-Diwali international match barely a week after the festival.

BCCI president Sourav Ganguly made it clear on Thursday that the match will not be cancelled at the last minute, but promised practical scheduling in future.

The cricketers have decided to shrug off the pollution threat and get on with the game. The poor air quality has left several Bangladesh players with sore throats and itchy eyes but coach Russell Domingo said they were focused on giving a good account of themselves on Sunday.

Domingo admitted that the prevailing smoggy conditions were not perfect, but said "no one is dying" as pollution is an issue in Bangladesh as well.

Several Bangladesh players and support staff wore face masks during Friday's training at the Kotla.

"Not obviously perfect with the smog but it's the same for both the teams," the South African said.

"It's not perfect, it's not ideal, but it's nothing we're going to complain about and moan over. We just got to get on with it... For sure we have some scratchy eyes, maybe a little sore throats now and then, but it's been okay. Nobody is being sick or dying or anything like that."

Delhi's notorious smog caused Sri Lankan bowler Suranga Lakmal to vomit on the field during a 2017 Test match at the Kotla and prompted questions about scheduling matches in northern India during winter.

Domingo referred to Bangladesh's own struggle with air pollution and said the team was not taking any special precaution.

"We know Sri Lanka struggled a bit last time. There's bit of pollution in Bangladesh as well. So it's not a massive shock to the system," Domingo said.

India's batting coach Vikram Rathour too played down the threat. "You're asking the wrong person I think," said the former player from Punjab.

"I've played all my life in north India. We're used to these conditions. We are not taking any special measure. The game has been scheduled and we're here to play."

A senior pulmonologist said cricketers, or other sportspersons, called on to play in conditions of severe air pollution could experience impaired performance.

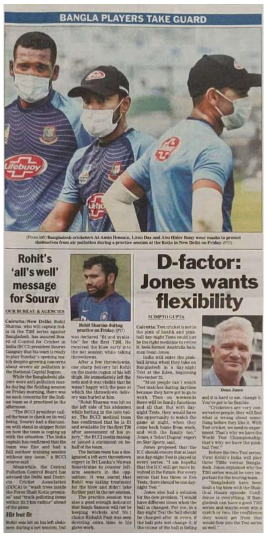
"Airflow requirements increase during periods of extreme pollution," Pratibha Gogia, a senior consultant respiratory medicine specialist in New Delhi, said.

High concentrations of pollutants could produce acute effects such as burning of eyes, cough or respiratory distress, she said.

It's not perfect, it's not ideal, but it's nothing we're going to complain about.... Nobody is being sick or dying or anything like that

RUSSELL DOMINGO
Bangladesh coach, on air quality in Delhi

(b) Segment



BANGLA PLAYERS TAKE GUARD

(From left) Bangladesh cricketers Al-Amin Hossain, Liton Das and Abu Hider Rony wear masks to protect themselves from air pollution during a practice session at the Kotla in New Delhi on Friday. (PTI)

Rohit's 'all's well' message for Sourav

OUR BUREAU & AGENCIES

Calcutta/New Delhi: Rohit Sharma, who will captain India in the T20I series against Bangladesh, has assured Board of Control for Cricket in India (BCCI) president Sourav Ganguly that his team is ready to play Sunday's opening match despite growing concerns about severe air pollution in the National Capital Region.

While the Bangladeshi players wore anti-pollution masks during the fielding session on Friday morning, there was no such concerns for the Indian team as it practised in the afternoon.

"The BCCI president called the team to check on its well being. Sourav had a discussion with stand in skipper Rohit on how the team is coping up with the situation. The India captain has confirmed that the team was fine and had a full outdoor training session without any issue," a BCCI source said.

Meanwhile, the Central Pollution Control Board has advised the Delhi and Districts Cricket Association (DDCA) to "wash trees inside the Feroz Shah Kotla premises" and "track polluting items within its 2 km radius" ahead of the game.

Hit but fit

Rohit was hit on his left abdomen during a net session, but was declared "fit and available" for the first T20I. He received the blow early into the net session while taking throwdowns.

Rohit Sharma during practice on Friday. (PTI)

After a few throwdowns, one sharp delivery hit Rohit on the inside region of his left thigh. He immediately left the nets and it was visible that he wasn't happy with the pace at which the throwdown delivery was hurled at him.

"Rohit Sharma was hit on the left side of his abdomen while batting in the nets today. The BCCI medical team has confirmed that he is fit and available for the first T20 after assessment of his injury," the BCCI media manager issued a statement on behalf of the team.

The Indian team has a designated a left-arm throwdown expert in Sri Lanka's Nuwan Seneviratne to counter left-arm seamers in the opposition. It was learnt that Rohit was taking treatment for the blow and didn't take further part in the net session.

The practice session was also a good enough indicator that Sanju Samson will not be keeping wickets and No.1 choice Rishabh Pant was seen devoting extra time to his glove work.

D-factor: Jones wants flexibility

SUDIPTO GUPTA

Calcutta: Test cricket is not in the pink of health and pink-ball day-night Tests could just be the right medicine to revive it, feels former Australia batsman Dean Jones.

India will enter the pink-ball arena when they take on Bangladesh in a day-night Test at the Eden, beginning November 22.

"Most people can't watch Test matches during daytime because they have got to go to work. Then on weekends there will be family functions and all that. But with day-night Tests, they would have the opportunity to watch the game at night, when they come back home from work, and with the family too," Jones, a 'Select Dugout' expert on Star Sports, said.

Jones proposed that the ICC should ensure that at least one day-night Test is played in every series. "I am hopeful that the ICC will get more involved in the future. For every series that has three or five Tests, there should be one day-night Test."

Jones also had a solution for the dew problem. "I would have different times when the ball is changed. For me, in a day-night Test the ball should be changed after 50 overs. If the ball gets wet change it. If the colour of the ball is fading and it is hard to see, change it. You've got to be flexible.

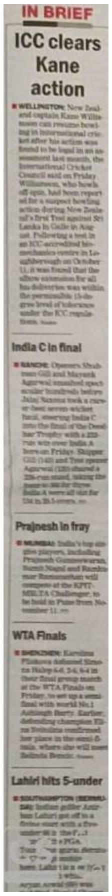
Dean Jones

"Cricketers are very conservative people, they will find what is wrong about something before they like it. With Test cricket, we need to experiment. That's why we have the World Test Championship, that's why we have the pink-ball Test."

Before the two-Test series, Virat Kohli's India will play three T20Is versus Bangladesh. Jones explained why the T20I series would be very important for the touring team.

"Bangladesh have been dealt a big blow with the Shakib-al Hasan episode. Confidence is everything. If Bangladesh can have a good T20I series and maybe even win a match or two, the confidence they would get from that would flow into the Test series as well."

(c) Segment



IN BRIEF

ICC clears Kane action

WELLINGTON: New Zealand captain Kane Williamson can resume bowling in international cricket after his action was found to be legal in an assessment last month, the International Cricket Council said on Friday. Williamson, who bowls off spin, had been reported for a suspect bowling action during New Zealand's first Test against Sri Lanka in Galle in August. Following a test in an ICC-accredited biomechanics centre in Loughborough on October 11, it was found that the elbow extension for all his deliveries was within the permissible 15-degree level of tolerance under the ICC regulations. Reuters

India C in final

RANCHI: Openers Shubman Gill and Mayank Agarwal smashed spectacular hundreds before Jalaj Saxena took a career-best seven-wicket haul, steering India C into the final of the Deodhar Trophy with a 232-run win over India A here on Friday. Skipper Gill (143) and Test opener Agarwal (120) shared a 226-run stand, taking the team to 366 for three. India A were all out for 134 in 29.5 overs. PTI

Prajnesh in fray

MUMBAI: India's top singles players, including Prajnesh Gunneswaran, Sumit Nagal and Ramkumar Ramanathan will compete at the KPIT-MSLTA Challenger, to be held in Pune from November 11. PTI

WTA Finals

SHENZHEN: Karolina Pliskova defeated Simona Halep 6-0, 2-6, 6-4 in their final group match at the WTA Finals on Friday, to set up a semi-final with world No.1 Ashleigh Barty. Earlier, defending champion Elina Svitolina confirmed her place in the semi-finals, where she will meet Belinda Bencic. Reuters

Lahiri hits 5-under

SOUTHAMPTON (BERMUDA): Indian golfer Anirban Lahiri got off to a fine start with a five-under 66 in the [illegible] PGA Tour [illegible] Lahiri [illegible] Aryan Arwal [illegible]

(d) Segment

D LIMITATIONS

D.1 Limitations

There are several limitations in the project. They can be listed as below:

Image Resolution: The input image must be of high resolution and large sized. If otherwise the segmentation will not take place.

Image Orientation: The Image must be in vertical orientation and must be perfectly straight.

Noise: The Image must be free of any noise or else erroneous outputs will come.

Nested Articles: In case of non-aligned nested articles the subimages must be again run to get all the articles separately.

D.2 Solution to the limitations

The limitations can be overcome by taking some measures like :

1. Removing noises by noise removal techniques.

2. Linking all the words of a same article by blackening the spaces between them. This will help identify even nested articles easily.

E FUTURE OF THE PROJECT

The project can be easily used as a starting platform for other projects like converting the articles into *.txt* files by using some *Optical Character Recognition* (OCR) tools. Then these *.txt* files can be segregated under different topics and genres.

Also this project will be a good platform to start other projects like Segregating Chapters of a Book.

F REFERENCE

[**1**] Wikipedia
https://www.wikipedia.org/.
[**2**] Stack Overflow
https://stackoverflow.com
[**3**] Overleaf
https://www.overleaf.com
[**4**] Image Processing in Java
Douglas A. Lyon